This is to certify that the
dissertation entitled

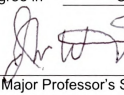
INCREASING THE CARBON FLOW INTO THE AROMATIC COMMON PATHWAY: BIOSYNTHESIS OF 3-DEHYDROSHIKIMIC ACID FROM D-GLUCOSE

presented by

Jian Yi

has been accepted towards fulfillment
of the requirements for the

Ph. D. degree in Organic Chemistry



Major Professor's Signature

2/27/04

Date

| DATE DUE | DATE DUE | DATE DUE |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

# INCREASING THE CARBON FLOW INTO THE AROMATIC COMMON PATHWAY: BIOSYNTHESIS OF 3-DEHYDROSHIKIMIC ACID FROM D-GLUCOSE

By

Jian Yi

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry

2004

## ABSTRACT

INCREASING THE CARBON FLOW INTO THE AROMATIC COMMON PATHWAY: BIOSYNTHESIS OF 3-DEHYDROSHIKIMIC ACID FROM D-GLUCOSE

By

Jian Yi

Metabolic engineering, DNA microarray and recombinant DNA technology were used to develop strategies to increase the carbon flow directed into the shikimate pathway. 3-Dehydroshikimic acid (DHS) is the most advanced intermediate shared by both biosynthesis of aromatic amino acids and biocatalytic syntheses of a variety of value-added chemicals. Strategies elaborated to increase the yield and concentration of DHS synthesized from glucose are thus applicable to the microbial synthesis of a wide range of chemicals.

Shikimate pathway product yields in microbial synthesis are ultimately limited by the glucose transport mechanism. To increase the availability of phosphoenolpyruvate (PEP), overexpression of PEP synthase or alteration of glucose transport by Glf-mediated facilitated diffusion or the GalP galactose permease in *E. coli* constructs led to increased yields of DHS and shikimate pathway byproducts. Overexpression of PEP synthase was currently leading to the synthesis of higher yields of DHS and shikimate pathway byproducts relative to alteration of glucose transport. In addition, the production of DHS can be enhanced by changing fermentor-controlled pH from 7 to 6.

DNA microarray technology was used to study gene expression profiles under fed-batch fermentor conditions. The results showed that a simple overexpression of PEP

synthase led to dramatic expression changes of many genes. Based on the microarray study, a more stable DAHP synthase ($AroG^{FBR}$) was isolated by PCR mutagenesis and has its advantage over previously used DAHP synthase ($AroF^{FBR}$) because of its stability over the course of a fermentation run.

The export of shikimic acid or DHS was also studied. Two approaches to identify the shikimate export protein were attempted without success. The first approach attempted to identify a mutant deficient in shikimate export while the second approach was based on the overexpression of a shikimate export protein. The failure of both approaches is consistent with the possibility that multiple proteins exist for the export of shikimic acid. Furthermore, the determination that intracellular shikimic acid concentrations are much lower than extracellular concentrations suggests that export of this hydroaromatic is not a limiting factor during microbial synthesis.

To my parents and my wife

For their love and support

## ACKNOWLEDGMENTS

First, I would like to thank my research adviser, Professor John W. Frost, for his patience, guidance and understanding during my time in the group. His excellent spoken and written skill, his aggressiveness, his pertaining to detail and hard working attitude will influence me for my entire life. In addition, I want to thank the members of my graduate committee, Professor Baker Gregory, Professor John McCracken, and Professor Geiger James for their discussion of the thesis.

I am grateful for Dr. Karen Draths for her friendship, patient guidance in the experimental techniques and intelligent input to my research. Her word " one complete experiment was better than 10 incomplete experiments" will be remembered forever. Meanwhile I am appreciative of current and former group members: Dr. Kai Li, Dr. Feng Tian, Dr. Spiros Kambourakis, Dr. Sunil Chandran, Dr. Jessica Barker, Dr. Dave Knop, Dr. Chad Hansen, Dr. Padmesh Venkitasubramanian, Dr. Dongming Xie, Dr. Jihane Achkar, Ningqing Ran, Wei Niu, Jiantao Guo, Heather Steuben, Mapitso Molefe, Xiaofei Jia, Wensheng Li, Jinsong Yang, Justas Jancauskas, Mankit Lau, Adam Bush and Lee Stephen Sing Kin, for their assistance and friendship. We are a big family!

This thesis is dedicated to my parents, Guangyou Yi and Liangying Yao, and my wife, Chunhui Li. Without their love and support, this thesis would never have happened.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| Ac | acetyl |
| ADP | adenosine diphosphate |
| ATP | adenosine triphosphate |
| Ap | ampicillin |
| $Ap^R$ | ampicillin resistance gene |
| bp | base pair |
| CA | chorismic acid |
| CIAP | calf intestinal alkaline phosphatase |
| Cm | chloramphenicol |
| $Cm^R$ | chloramphenicol resistance gene |
| COMT | catechol-*O*-methyltransferase |
| DAHP | 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate |
| DCU | digital control unit |
| DEAE | diethylaminoethyl |
| DHQ | 3-dehydroquinic acid |
| DHS | 3-dehydroshikimic acid |
| DO | dissolved oxygen |
| DTT | dithiothreitol |
| E4P | D-erythrose 4-phosphate |
| EMP | Embden-Meyerhof pathway |
| EPSP | 5-enolpyruvoylshikimate 3-phosphate |
| FBR | feedback resistant |
| GA | gallic acid |

| | |
|---|---|
| h | hour |
| HPLC | high pressure liquid chromatography |
| IPTG | isopropyl β-D-thiogalactopyranoside |
| Kan | kanamycin |
| *Kan*$^R$ | kanamycin resistance gene |
| kb | kilobase |
| kg | kilogram |
| Km | Michaelis constant |
| LB | luria broth |
| M | molar |
| M9 | minimal salts |
| min | minute |
| mL | milliliter |
| μL | microliter |
| mM | millimolar |
| μM | micromolar |
| mRNA | messenger RNA |
| NAD | nicotinamide adenine dinucleotide, oxidized form |
| NADH | nicotinamide adenine dinucleotide, reduced form |
| NADP | nicotinamide adenine dinucleotide phosphate, oxidized form |
| NADPH | nicotinamide adenine dinucleotide phosphate, reduced form |
| NMR | nuclear magnetic resonance spectroscopy |
| OD | optical density |
| ORF | open reading frame |
| PCA | protocatechuic acid |
| PEG | polyethylene glycol |

| | |
|---|---|
| PEP | phosphoenolpyruvic acid |
| pfu | plaque forming units |
| PHB | *p*-hydroxybenzoic acid |
| PID | proportional-integral-derivative |
| PCR | polymerase chain reaction |
| Phe | L-phenylalanine |
| psi | pounds per square inch |
| PTS | phosphotransferase system |
| QA | quinic acid |
| rpm | rotations per minute |
| SA | shikimic acid |
| SAM | *S*-adenosylmethionine |
| SDS | sodium dodecyl sulfate |
| S3P | shikimate 3-phosphate |
| Tc | tetracycline |
| TCA | tricarboxylic acid |
| Trp | L-tryptophan |
| TSP | sodium 3-(trimethylsilyl)propionate-*2,2,3,3*-$d_4$ |
| Tyr | L-tyrosine |
| UV | ultraviolet |

# CHAPTER 1

## Introduction

Chemistry is moving into an era in which renewable resources and starting materials such as D-glucose, D-xylose, and L-arabinose will play an important role in industrial chemical manufacture.[1] Current chemical manufacture typically relies on abiotic, chemical catalysts and on starting materials derived from petroleum. As a nonrenewable natural resource, petroleum has several negative environmental and geopolitical problems associated with its use. For example, aromatics are currently derived predominantly from the benzene, toluene, xylene (BTX) fraction of petroleum refining. With annual production levels at 1.9 billion gallons in the United States, benzene is the most important component of the BTX fraction.[2] Benzene is a potent carcinogen,[3] and it is also included by the Environmental Protection Agency on the list of chemicals covered by the Chemical Manufacturing Rule that requires drastic reductions in emissions of hazardous organic air pollutants.[4]

Beyond the health costs associated with benzene, the costs of deriving this starting material from nonrenewable fossil fuel feedstocks are important to consider.[5] Oil spills and land reclamation along with geopolitical complications substantially add to the true cost of aromatic manufacture from fossil fuel-derived BTX. By contrast, plant-derived starch, cellulose, and hemicellulose are abundant, renewable sources of glucose, xylose, and arabinose, which can be used as carbon source for microbial synthesis of a variety of chemicals.[6] The avoidance of toxic intermediates, reagents, and byproducts, and the ability to synthesize a chemical by microbial biocatalysis from renewable feedstocks and nontoxic starting materials presents itself as an appealing alternative to traditional chemical manufacture.[6]

In chapter 2 of this thesis, two basic strategies were examined to surmount the fundamental limitation on the yields of 3-dehydroshikimic acid and shikimic acid imposed by PTS-mediated glucose transport. In chapter 3 of this thesis, DNA microarray technology was applied to detect differential transcription profiles of relative "high-yield"

cells and relative "low-yield" cells. The gene targets for further yield improvement for biocatalyzed synthesis of 3-dehydroshikimic acid were identified. In Chapter 4 of this thesis, efforts towards identification of the shikimic acid efflux protein is reported and whether efflux of shikimic acid is a limiting factor during microbial synthesis of shikimic acid is established. All the chapters focus on developing strategies to increase the carbon flow directed into the shikimate pathway to improve titers and yields of biosynthesized chemicals. 3-Dehydroshikimic acid (DHS) is the most advanced intermediate shared by both biosynthesis of aromatic amino acids (Figure 1) and biocatalytic syntheses of a variety of value-added chemicals. Strategies elaborated to increase the yield and concentration of 3-dehydroshikimic acid synthesized from glucose are thus applicable to the microbial synthesis of a wide range of chemicals.

## The shikimate pathway

The shikimate pathway, also referred to as the common pathway of aromatic amino acid biosynthesis, has been the subject of considerable study due to its role in the transformation of simple carbohydrate precursors into aromatics in plants, bacteria, fungi, and molds.[7] It consists of seven enzyme-catalyzed reactions converting phosphoenolpyruvic acid (PEP) and erythrose 4-phosphate (E4P) into chorismic acid (Figure 1). Three terminal pathways then lead from chorismic acid to L-tryptophan, L-tyrosine, and L-phenylalanine. In addition, biosynthetic pathways leading to ubiquinone, folic acid, enterochelin and huge number of secondary metabolites also branch away from the common pathway at chorismic acid (Figure 1).[5] Folic acid-derived coenzymes are frequently involved in the biosynthetic transfer of one carbon fragment, ubiquinones are involved in electron transport, and enterochelin is an iron chelator responsible for iron uptake in numerous microorganisms.

Individual pathway enzymes have received attention due to the novel mechanisms

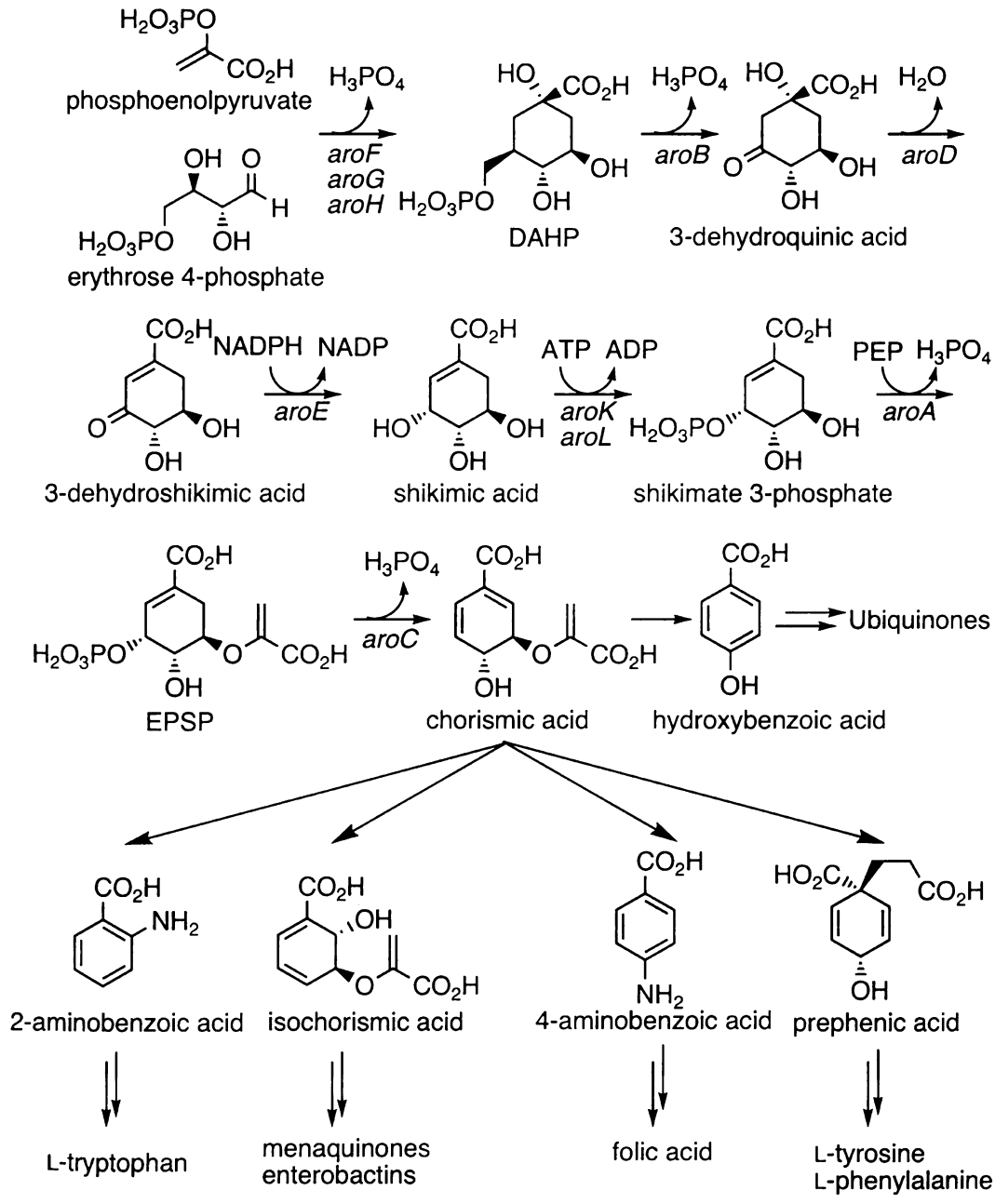



**Figure 1. The common pathway of aromatic amino acid biosynthesis and pathways beyond chorismate.** Intermediates (abbreviations): 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate (DAHP), 3-dehydroquinic acid (DHQ), 5-enolpyruvyl-shikimate-3-phosphate (EPSP). Enzymes (genes): *aroF,* DAHP synthase (Tyrosine); *aroG*, DAHP synthase (phenylalanine); *aroH*, DAHP synthase (Tryptophan); *aroB*, DHQ synthase; *aroD*, DHQ dehydratase; *aroE*, shikimate dehydrogenase; *aroL*, shikimate kinase; *aroK*, shikimate kinase; *aroA*, EPSP synthase; *aroC*, chorismate synthase.

employed during turnover of substrate into product and as targets for inhibition. The existence of the shikimate pathway in plants and bacteria but not in humans provides an appealing mode of action for enzyme-targeted herbicides and antibiotics.

The substrates and products of the seven enzymatic reactions that convert PEP and E4P into chorismic acid (Figure 1) were identified by the early 1960s from studies of bacterial auxotrophs of *Escherichia coli* and *Klebsiella aerogenes*. The first committed step of aromatic amino acid biosynthesis involves the condensation of PEP and E4P to form 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate (DAHP) and is catalyzed by DAHP synthase.[5] Three isozymes of DAHP synthase exist in *E. coli*, each of which is sensitive to feedback inhibition by one of the three aromatic amino acids. The genes *aroF*, *aroG* and *aroH* encode for tyrosine-sensitive, phenylalanine-sensitive, and tryptophan-sensitive isozymes of DAHP synthase, respectively. DAHP is converted into 3-dehydroquinic acid (DHQ) by DHQ synthase which is encoded by *aroB*[8] in a complex reaction where $NAD^+$ is a catalytic requirement.[9] A *syn* elimination of water from DHQ affords 3-dehydroshikimic acid (DHS).[10] This reaction is catalyzed by DHQ dehydratase, which is encoded by *aroD*. Reduction of DHS to shikimic acid in the presence of NADPH is catalyzed by *aroE*-encoded shikimate dehydrogenase.[11] Shikimic acid is further converted to shikimate 3-phosphate by phosphoryl group transfer from ATP. This reaction is catalyzed by two isozymes of shikimate kinase encoded by the loci *aroL*[12] and *aroK*.[13] 5-Enolpyruvoylshikimate 3-phosphate (EPSP) synthase, encoded by *aroA*,[14] catalyzes the reversible condensation of shikimate 3-phosphate and PEP. Product EPSP is formed along with inorganic phosphate. The last enzyme of the common pathway is chorismate synthase.[15] Encoded by *aroC*, it catalyzes the elimination of inorganic phosphate from EPSP to afford chorismic acid.

## Microbial syntheses of value-added chemicals utilizing the shikimate pathway

### L-Phenylalanine, L-tryptophan and L-tyrosine

Aromatic amino acids such as L-phenylalanine, L-tryptophan and L-tyrosine are prominently among the chemicals being microbially manufactured from glucose.[6, 16] L-Tryptophan (market volume 1.1 – 1.2 x $10^7$ kg/year[17]) is produced predominantly for use as a feed additive. Introduction of naphthalene dioxygenase into a tryptophan-synthesizing microbe that also expresses tryptophanase results in biocatalytic synthesis of indigo (Figure 2),[18] the vat dye that gives blue jeans their faded-blue coloration. L-Phenylalanine (market volume 5 – 6 x $10^5$ kg/year[17]) is produced predominantly for the production of the low-calorie sweetener aspartame (Figure 2).[16] Other uses for L-phenylalanine are its use in infusion fluids, in food additives, as intermediate for synthesis of pharmaceuticals (HIV protease inhibitor, anti-inflammatory drugs, rennin inhibitors, etc.) and as a flavor enhancer.[6b] L-tyrosine (market volume 1.5 x $10^5$ kg/year[17]) is produced at a small scale and its use for the production of the anti-Parkinson's drug 3,4-dihydroxyphenyl L-alanine (L-DOPA)[19] (Figure 2), for the treatment of Basedow's disease and as a dietary supplement.[6b]

Microbial production of L-phenylalanine has been focused mainly on *E. coli*, *C. glutamicum* and *Brevibacterium* strains.[20] Classic methods have been applied to screen for auxotrophs and mutants with feedback deregulated key enzymes. Sugimoto *et al.* cloned the genes encoding feedback resistant forms of DAHP synthase, $aroF^{FBR}$, and chorismate mutase/prephenate dehydratase, $pheA^{FBR}$, into a temperature-controllable expression vector. An L-tyrosine auxotrophic *E. coli* strain carrying this plasmid can reach a titer of 46 g/L and a productivity of 0.85 g/L/h.[21] To obtain L-tyrosine overproducers, most attention has been focused on screening for regulatory and auxotrophic mutants. These were mostly strains of *E. coli*, *Bacillus subtilis* or various coryneform bacteria.[22] By application of recombinant DNA technology additional improvements of these L-tyrosine producing strains have been

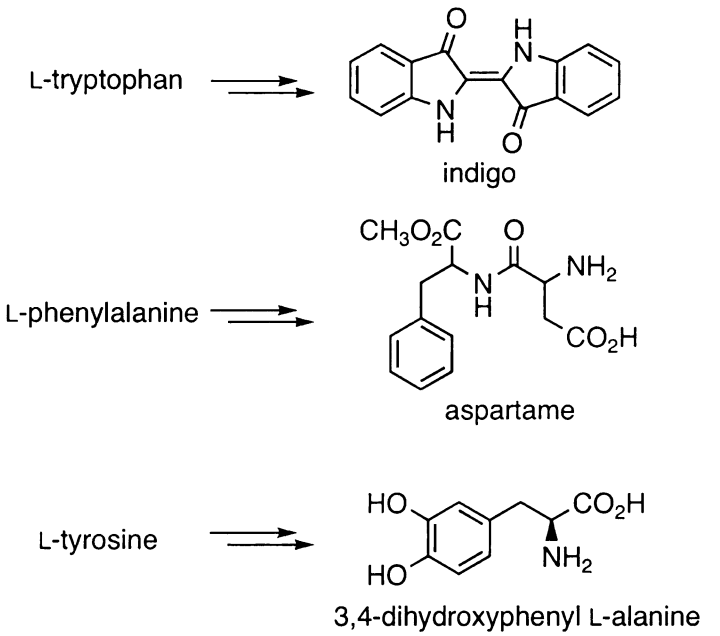



**Figure 2. Chemicals synthesized from tryptophan, tyrosine, and phenylalanine.**

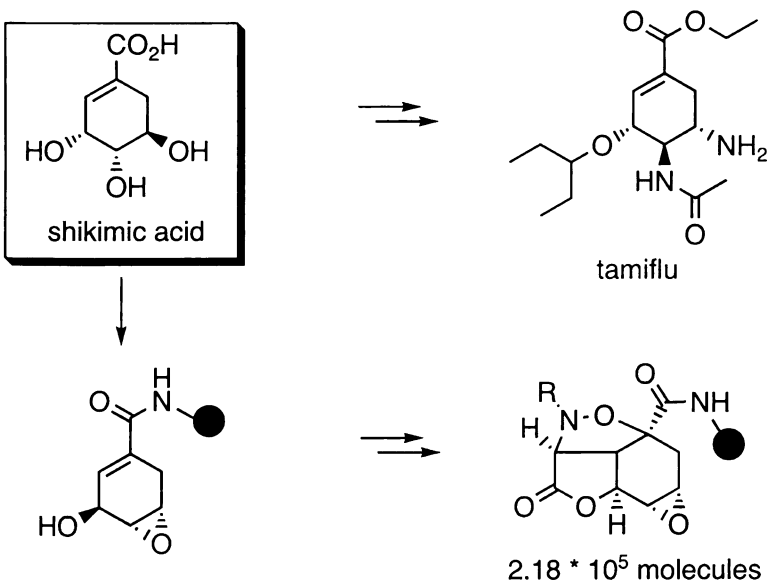



**Figure 3. Shikimic acid's use in synthetic chemistry.**

made.[23] Metabolic engineering for production of L-tryptophan, has primarily been carried out in *E. coli* (reviewed by Berry[24]), *Corynebacterium*[25] and *Bacillus*.[26] Alterations in the central carbon metabolism and the common aromatic amino acid pathway, including its

regulation, were carried out. In addition, overexpression of the genes encoding the tryptophan branch of the aromatic amino acid pathway was performed.[24]

**Shikimic acid**

Shikimic acid has been exploited by Schreiber and coworkers in their use of this hydroaromatic as the core scaffold for the synthesis of large combinatorial libraries of molecules (Figure 3).[27] In addition, shikimic acid is a valuable chiral synthon used in the synthesis of tamiflu (soeltaminvir, Figure 3),[28] which is a potent inhibitor of both influenza A and influenza B neuraminidases. Tamiflu was approved by the FDA and is marketed by Roche as an orally administered antiinfluenza agent. Tamiflu's six-membered carbocyclic ring, carboxylic acid, and backbone stereocenters have led to the use of shikimic acid as the starting material. Nonetheless, shikimic acid's use in synthetic chemistry has been restricted by its limited availability and prohibitive cost. Shikimic acid is isolated from plant sources. However, isolation of shikimic acid from the pericarps of *Illicium* is a tedious procedure.[29] The supply of *Illicium* does not benefit from monoculture.[30] When Tamiflu came out of clinical trials, Roche was confronted by a price for shikimic acid that precluded its viable use as a pharma starting material.

To alleviate supply issues, *E. coli* SP1.1/pKD12.138 was constructed for microbial conversion of glucose into shikimic acid.[31] Shikimic acid production using *E. coli* SP1.1/pKD12.138 has been scaled up to a 50,000 L process and is used by Roche as one of its sources of shikimic acid. This shikimate-synthesizing biocatalyst resulted from disruption of the genomic *aroL* and *aroK* loci (Figure 1) in *E. coli* and overexpression of feedback-insensitive, $aroF^{FBR}$-encoded DAHP synthase to channel more carbon into the common pathway.[32] Another advantage of a microbe-catalyzed conversion of glucose into shikimic acid is the ability to make tons of shikimic acid from glucose within a 48 h time frame. This abundant, reliable source of shikimic acid contrasts with the need to wait for an

entire growing season for *Illicium* plants to mature and circumvents the problems of adverse weather destroying production for an entire *Illicium* growing season.

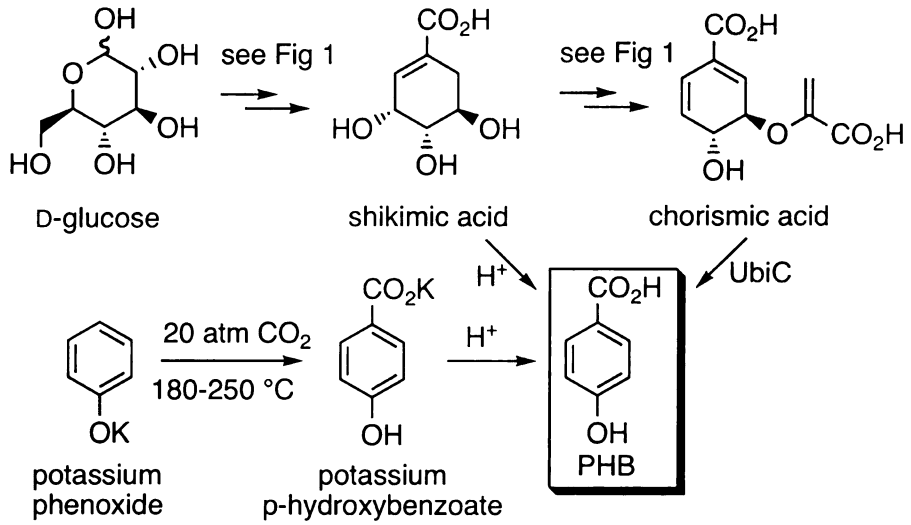



**Figure 4. Comparison of *p*-hydroxybenzoic acid (PHB) synthesis from glucose versus potassuim phenoxide.** Enzyme: Chorismate-pyruvate lyase (UbiC).

***p*-Hydroxybenzoic acid**

*p*-Hydroxybenzoic acid is a component of liquid crystal polymers such as Xydar,[33] which have attracted considerable attention because of their use in high-performance applications. Esters of *p*-hydroxybenzoic acid are also widely used as food preservatives.[34] *p*-Hydroxybenzoic acid is currently manufactured by Kolbe-Schmitt reaction of dried potassium phenoxide with 20 atm dry carbon dioxide at 180-250 °C (Figure 4).[34] Product *p*-hydroxybenzoate potassium salt is converted to its free acid upon addition of mineral acid. Besides the required temperatures and pressures, *p*-hydroxybenzoic acid manufacture has to contend with handling of phenol which is listed as a highly toxic, corrosive chemical.[35]

Chorismic acid can also be converted directly into *p*-hydroxybenzoic acid in a reaction catalyzed by *ubiC*-encoded[36] chorismate-pyruvate lyase (Figure 4). An *E. coli* biocatalyst has been constructed consisting of a genome-localized *aroAaroLaroBaroCkan*$^R$

cassette and plasmid-localized *aroF*$^{FBR}$, *tktA*, and *ubiC*. The genes encoding o chorismate-utilizing enzymes are disrupted to prevent biocatalytic conversion of choris acid into anthranilic acid and prephenic acid (Figure 1). The biocatalyst can synthesize g/L of *p*-hydroxybenzoic acid from glucose under fed-batch fermentor conditions.[37] Hydroxybenzoic acid can also be synthesized by means of chemical dehydration shikimic acid. This reaction is catalyzed by 1 M sulfuric acid in refluxing acetic a (Figure 4).[37]

**Quinic acid and hydroquinone**

Quinic acid is another useful molecule that can be microbially synthesized thro the utilization of the shikimate pathway (Figure 5).[38] Quinic acid's six-membe carbocyclic ring arrayed with stereogenic centers has been extensively used as a ch synthon and starting material in natural product synthesis.[39] It is currently obtained by expensive isolation from plant sources. An *E. coli* biocatalyst has been constructed mutational inactivation of *aroD*-encoded DHQ dehydratase, overexpression of *aroF* encoded feedback-insensitive DAHP synthase and overexpression of *aroE*-enco shikimate dehydrogenase. The biocatalyst can synthesize 49 g/L of quinic acid fr glucose in 20% (mol/mol) yield.[32] Quinic acid is then chemically converted hydroquinone and benzoquione by simple chemical oxidation (Figure 6).[38] Other hi yielding chemical methodology has also been elaborated for conversion of quinic aci these fermentation broths into hydroquinone.[40] Oxidative decarboxylation of quinic aci clarified, decolorized, ammonium ion-free fermentation broth with NaOCl and subsequ dehydration of the intermediate afforded purified hydroquinone in 87% yield. Halide-f oxidative decarboxylation of quinic acid in fermentation broth with stoichiometric quant of $(NH_4)_2Ce(SO_4)_3$ and $V_2O_5$ afforded hydroquinone in 91% and 85% yield, respectiv Conditions suitable for oxidative decarboxylation of quinic acid with catalytic amount metal oxidants were also identified. $Ag_3PO_4$ at 2 mol % relative to quinic aci

fermentation broth catalyzed the formation of hydroquinone in 74% yield with $K_2S_2$ serving as the cooxidant. Selective reduction of photoactivated silver ion with hydroquino is the basis for this organic's widespread use in photography. Benzoquinone is important chemical precursor in the manufacture of various chemicals.[41]

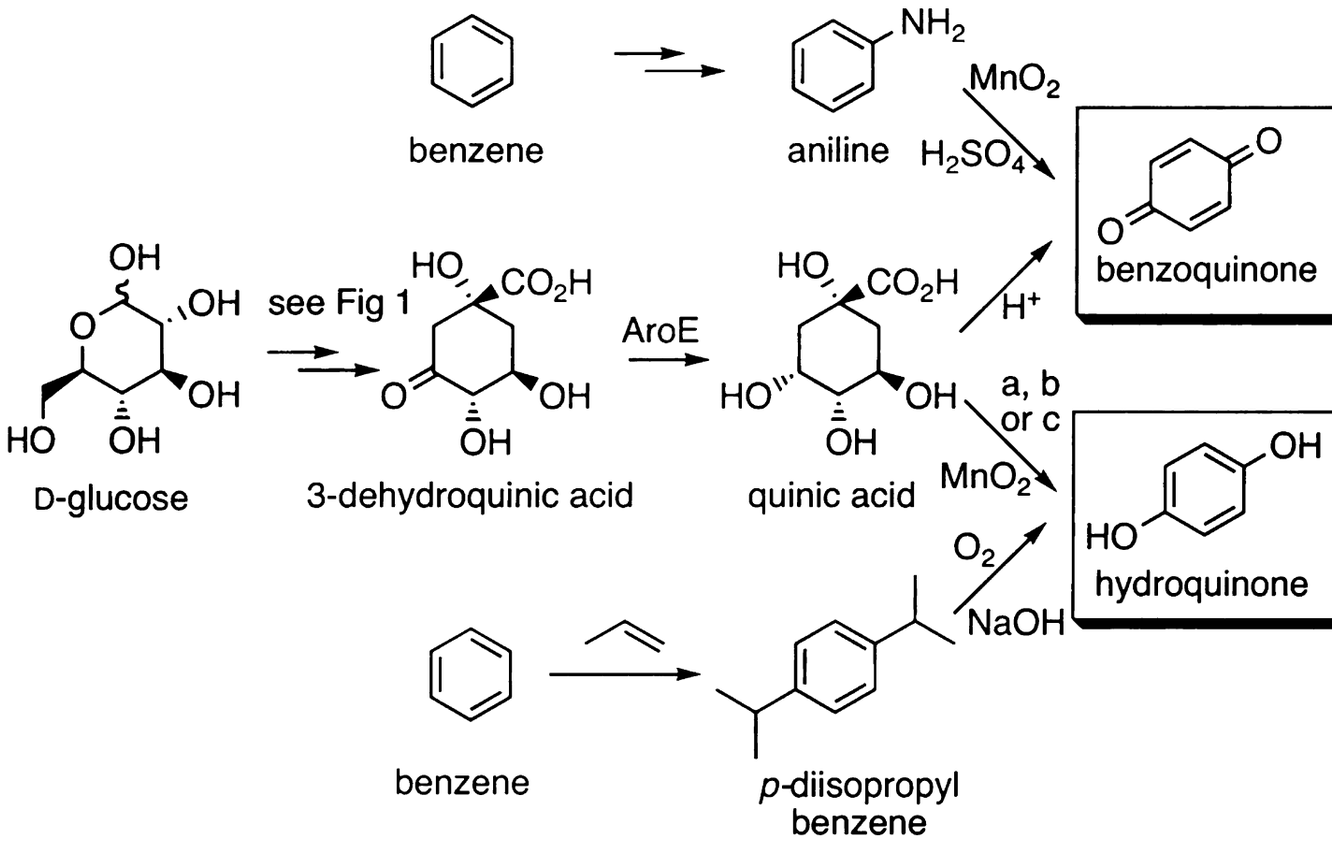



**Figure 5. Comparison of benzoquinone and hydroquinone synthesis from gluco and benzene.** Enzyme: Shikimate dehydrogenase (AroE). Key: a) i. NaOCl, isopropanol, iii. reflux; b) $(NH_4)_2Ce(SO_4)_3$ (2.4 equiv), 30 min; reflux, 10 h. c) $Ag_3PO_4$ ( mol %), $K_2S_2O_8$, 50 °C, ii.reflux.

Hydroquinone is produced globally at volumes of 4.5 - 5.0 x $10^7$ kg/year.[42] T dominant route to manufacture hydroquinone is through hydroperoxidative synthesis The *p*-diisopropylbenzene is synthesized by zeolite-catalyzed Friedel-Crafts reaction benzene or cumene with propylene or isopropanol. Air oxidation of the diisopropylbenzene proceeds at 90 - 100 °C in an aqueous NaOH solution also containi organic bases along with cobalt or copper salts. Hydroperoxycarbinol and dicarbinol produced along with the dihydroperoxide during air oxidation. Treatment with acid a $H_2O_2$ converts the hydroperoxycarbinol and dicarbonyl to the dihydroperoxide which

cleaved to form acetone and hydroquinone. During the acidic cleavage, explosive orga peroxide can form which presents a safety hazard.[42] Benzoquinone is prima manufactured from oxidation of aniline (Figure 5). The oxidant employed is $MnO_2$ aqueous $H_2SO_4$ (Figure 5). Benzoquinone can also be reduced by Fe or hydrogenate afford hydroquinone. Although accounting for approximately 10% of hydroquin production,[42] this manufacturing route generates large quantities of $MnSO_4$, $(NH_4)_2S$ and iron oxide salts.[42]

**3-Dehydroshikimic acid**

3-Dehydroshikimic acid (DHS), a powerful antioxidant,[43] is an intermediate in biosynthesis of aromatic natural products derived from chorismic acid and is also a bra point from which aromatic natural products can be biosynthesized without intermediacy chorismic acid (Figure 6). The aromatization of 3-dehydroshikimic acid to protocatech acid is crucial to the value of 3-dehydroshikimic acid as an intermediate towards synthesis of several commodity and fine chemicals. Protocatechuic acid is a branch poin the biosynthesis of catechol, *cis, cis*-muconic acid, vanillin, gallic acid and pyrogal Hydrogenation of *cis-cis*-muconic acid affords adipic acid. The Frost group developed *coli* KL3/pKL5.17A as a construct capable of converting glucose into 3-dehydroshiki acid.[44] This 3-dehydroshikimate-synthesizing biocatalyst resulted from disruption of genomic *aroE* loci (Figure 1) in *E. coli* and overexpression of feedback-insensiti $aroF^{FBR}$-encoded DAHP synthase and transketolase to channel more carbon into common pathway.[44] Progress made in improving the synthesis of 3-dehydroshikimic a can be viewed as progress made towards the synthesis of increased yields of shikim pathway products biosynthesized either with or without the intermediacy of chorismic aci

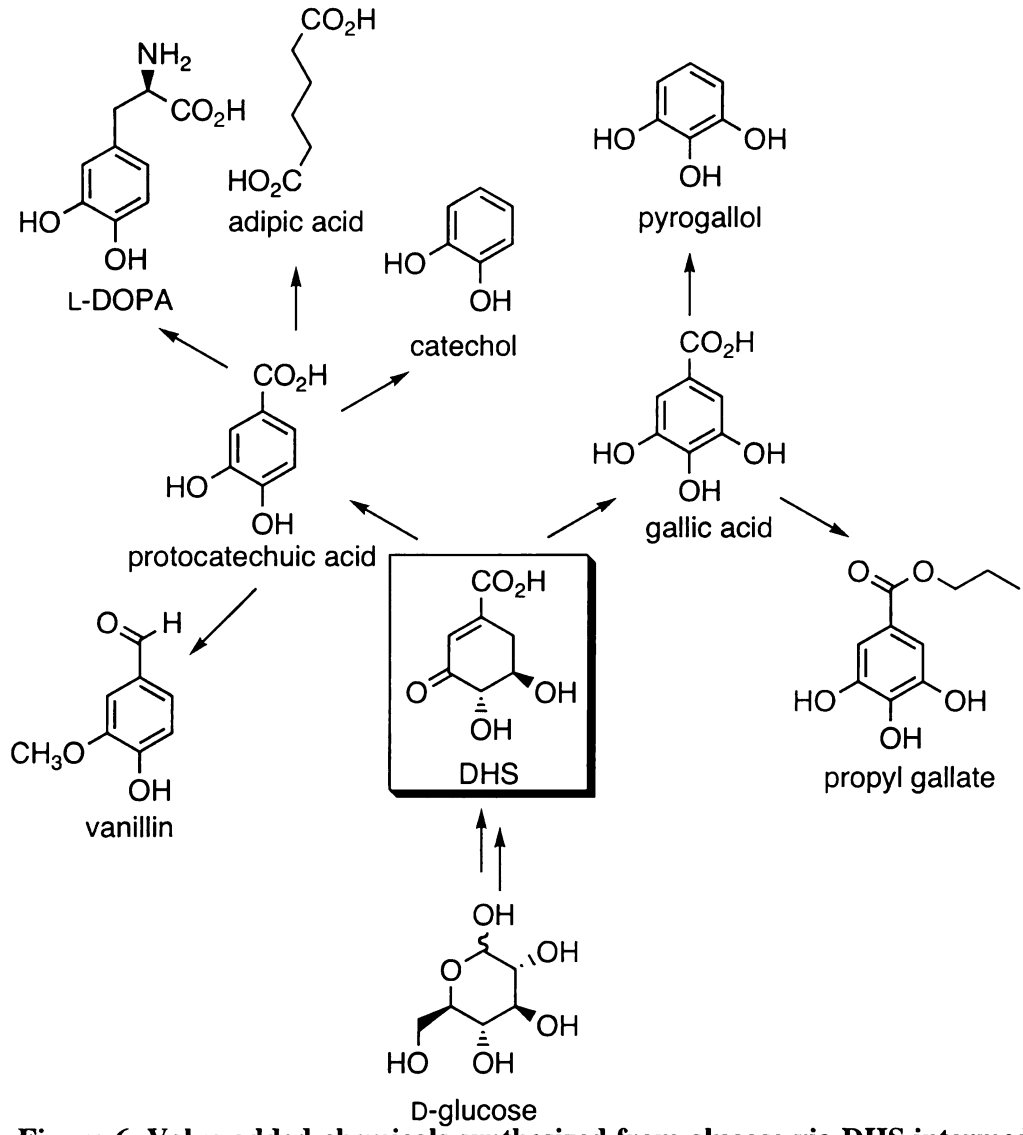



**Figure 6. Value added-chemicals synthesized from glucose via DHS intermediacy.**

**Gallic acid and pyrogallol**

Gallic acid can be isolated from gall nuts or from tara powder made from ground seed pods of the Peruvian tree species *Coulteria tinctoria*. Although gallic acid is available from natural sources, an inherent danger is its undependable supply. Neither the gall nuts acquired from insect carapaces nor the tara powder obtained from ground seed pods are cultivated in a controlled fashion. Both sources are highly susceptible to drought or flood conditions. Chemical oxidation of DHS can afford gallic acid.[45] DHS can also be converted to gallic acid biocatalytically via protocatechuic acid (PCA) intermediary (Figure

7). DHS dehydratase, which is encoded by the *aroZ* locus isolated from *Klebsiella pneumoniae*, converts DHS to PCA. Hydroxylation of PCA catalyzed by a mutant isozyme of *p*-hydroxybenzoate hydroxylase (*pobA**) isolated from *Pseudomonas fluorescens* then affords gallic acid. An *E. coli* biocatalyst has been constructed which synthesizes approximately 20 g/L of gallic acid from glucose under fed-batch fermentation conditions.[45]

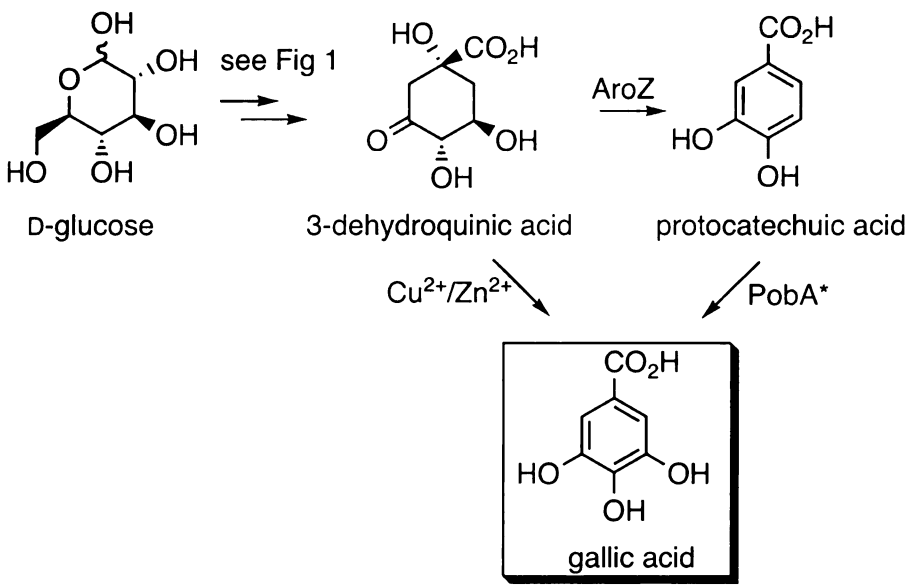



**Figure 7. Synthesis of gallic acid from glucose.** Enzyme: DHS dehydratase (AroZ), *p*-hydroxybenzoate hydroxylase mutant (*pobA**).

Derivatives of gallic acid include: propyl gallate (Figure 6), an important food-grade antioxidant; trimethoprim, an antibacterial agent; and pyrogallol, a product of chemical or enzymatic decarboxylation of gallic acid. Pyrogallol is used in photographic developing solutions and is a key building block in the manufacture of the insecticide bendiocarb.[46] Decarboxylation of gallic acid to pyrogallol was conducted by the construct *E. coli* RB791*serA*::*aroB*/pSK6.234 expressing a copy of *aroY*-encoded protocatechuic acid decarboxylase. Addition of gallic acid to a batch culture of *E. coli* RB791*serA*::*aroB*/pSK6.234 at stationary phase and at 33 °C and neutral pH afforded pyrogallol in 97% yield.[45] The toxicity of pyrogallol towards growing *E. coli* cells

precluded the direct synthesis of pyrogallol from D-glucose using a single microbi construct.[45]

## Catechol

Global, non-captive production of catechol is approximately 2.5 x $10^4$ tons pe year.[47] Estimating total world production is difficult since catechol is often part of captiv markets where it is produced and then converted into higher, value-added products befor ever seeing the open market as catechol. Chemical products derived from catechol includ pharmaceuticals (L-DOPA, adrenaline, papaverine), flavors (vanillin, eugenol, isoeugenol agrochemicals (carbofuran, propoxur), and polymerization inhibitors and antioxidants (4 *tert*-butylcatechol, veratrol).[48] Most catechol production begins with Friedel-Craft alkylation of benzene to afford cumene (Figure 8). Subsequent Hock-type, air oxidation c the cumene leads to formation of acetone and phenol. The phenol is then oxidized to mixture of catechol and hydroquinone using 70% hydrogen peroxide either in the presenc of transition metal catalysts or in formic acid solution where performic acid is the actua oxidant. Catechol and hydroquinone are separated by distillation (Figure 8).[48]

Catechol is also microbially synthesized from DHS (Figure 8).[49] A geneticall modified *E. coli* strain mediates the dehydration of DHS to form protocatechuic aci (Figure 8), which then undergoes an enzyme-catalyzed, non-oxidative decarboxylation t catechol. As in biocatalytic synthesis of gallic acid, dehydration of DHS to protocatechui acid is catalyzed by *aroZ*-encoded DHS dehydratase. Conversion of protocatechuic aci into catechol is catalyzed by *aroY*-encoded PCA decarboxylase, which was isolated from *K pneumoniae*. Although DHS and protocatechuic acid are intermediates in the conversion o glucose into catechol, only catechol accumulates in the culture supernatant. [49] However, th toxicity of catechol towards growing *E. coli* cells precluded the direct synthesis of catecho from D-glucose using a single microbial construct.[49]

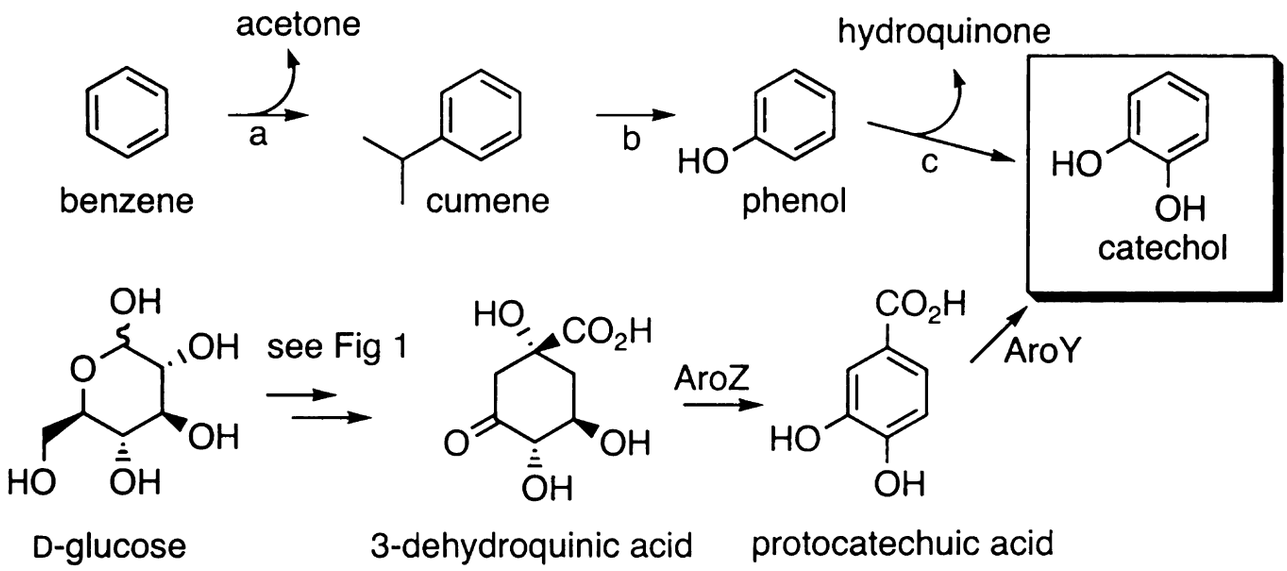



**Figure 8. Comparison of microbial and chemical synthesis of catechol.** Enzyme DHS dehydratase (AroZ), PCA decarboxylase (AroY). Key: a) propylene, solid $H_3PO_4$ catalyst, 200-260 °C, 400-600 psi.; b) $O_2$, 80-130 °C then $SO_2$, 60-100 °C; c) 70% $H_2O_2$ EDTA, $Fe^{2+}$ or $Co^{2+}$, 70-80 °C.

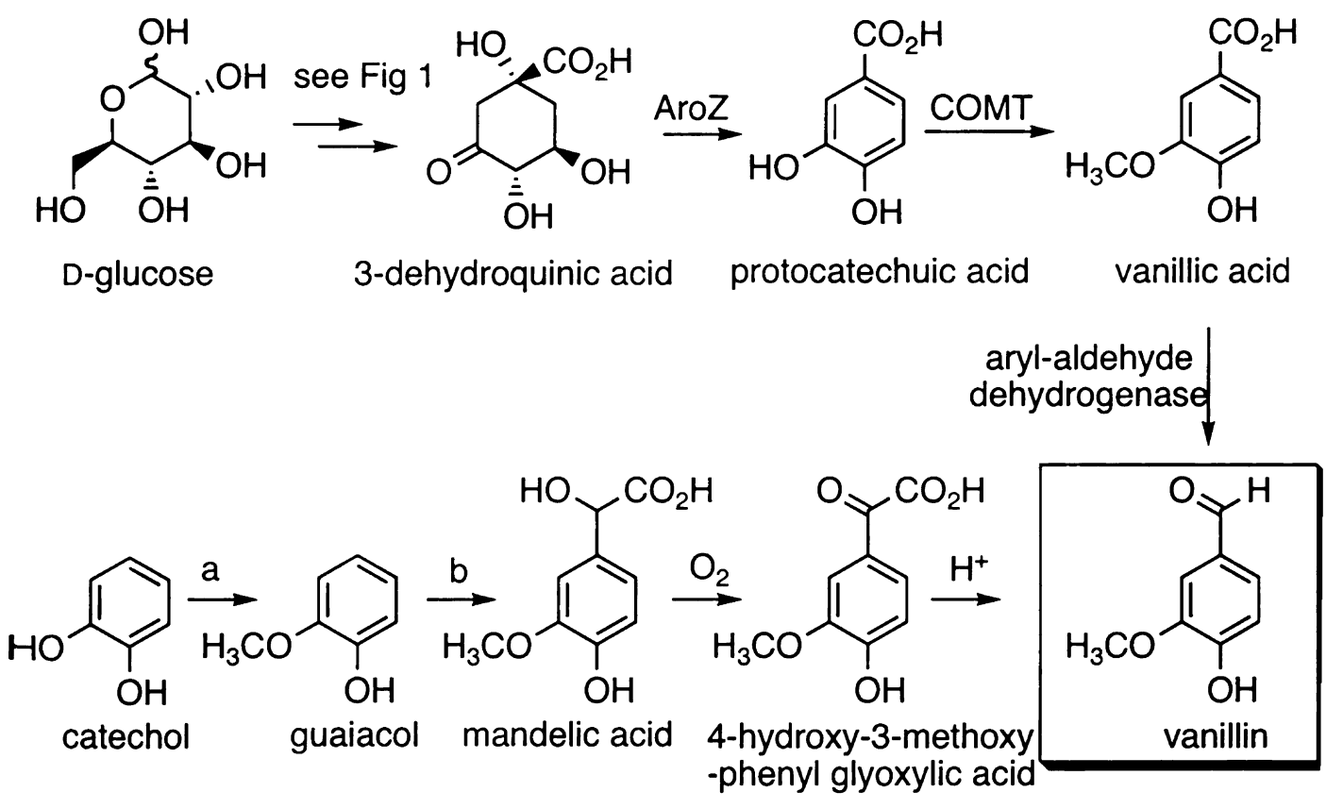



**Figure 9. Comparison of microbial and chemical synthesis of vanillin.** Enzymes DHS dehydratase (AroZ), catechol-*O*-methyltransferase (COMT). Key: a) $(CH_3O)_2SO_2$ b) glyoxylic acid.

## Vanillin

Vanillin (4-hydroxy-3-methoxybenzaldehyde; Figure 9) is the major component c natural vanilla, which is one of the most widely used and important flavouring material

worldwide. The source of vanilla is the bean, or pod, of the tropical *Vanilla* orchi (principally *Vanilla planifolia* Andrews, syn. *V. fragrans* (Salisb. Ames)).[50] Vanillin i fact occurs in trace amounts in other plants, including commercial products such a tobacco[51]; however, the pods of the *Vanilla* orchid still remain the only commercial sourc of natural vanillin. Although more than 12,000 tons of vanillin are produced each year, les than 1% of this is natural vanillin form *Vanilla*; the remainder is synthesized much mor cheaply via chemical processes.[50]

One synthetic route to produce vanillin begins with the condensation of guaiaco with glyoxylic acid (Figure 9). Air oxidation of the resulting mandelic acid afford phenylglyoxylic acid. Crude vanillin is obtained by acidification and decarboxylation of 4 hydroxyl-3-methoxyphenyl glyoxylic acid solution. Commercial grades are obtained b vacuum distillation and subsequent recrystallization.[52] Although this process is currentl the major source for synthetic vanillin, it has several inherent problems. Guaiacol ha historically been obtained from condensation of dimethyl sulfate with catechol. Dimethy sulfate is classified as a highly toxic, cancer-suspect agent. Catechol is listed as being toxi and corrosive while guaiacol is a toxic irritant. Catechol is currently manufactured fron benzene (Figure 9).

Synthesis of vanillin from glucose has been achieved using a microbe-catalyze conversion of glucose into vanillic acid followed by an enzyme-catalyzed reduction of th vanillic acid to afford vanillin (Figure 9).[53] A genetically engineered *E. coli* synthesize 5.0 g/L of vanillic acid from glucose under fed-batch fermentor conditions.[53] Aryl-aldehyd dehydrogenase purified from *Neurospora crassa* was used to reduce vanillic acid t vanillin in 66% isolated yield.[53]

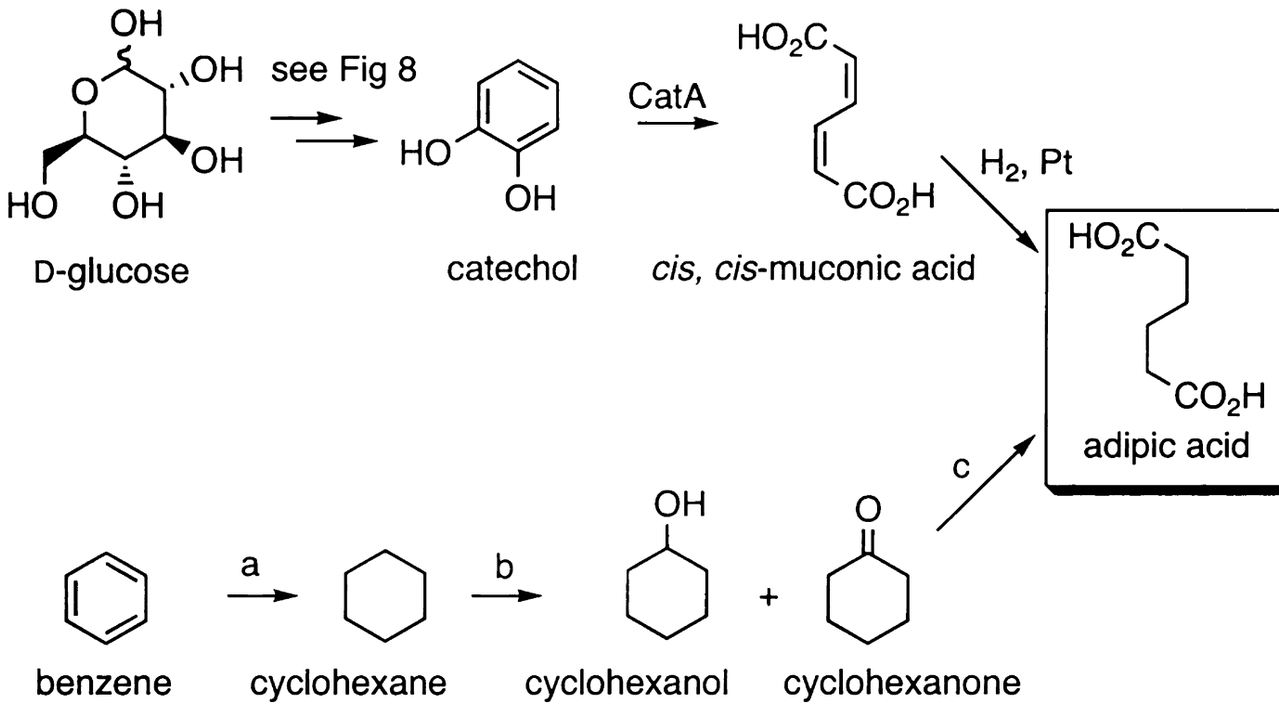



**Figure 10. Comparison of microbial and chemical synthesis of adipic acid.** Enz
CatA, catechol 1,2-dioxygenase. Key: a) Ni-$Al_2O_3$, $H_2$, 370-800 psi, 150-250 °C; b) Co
120-140 psi., 150-160 °C; c) Cu, $NH_4VO_3$, 60% $HNO_3$, 60-80 °C.

## Adipic acid

Adipic acid is primarily used for production of nylon-6,6. The annual g
demand for nylon exceeds 4.4 x $10^9$ kg.[54] Benzene is the principal starting material
which adipic acid is currently synthesized. Hydrogenation of benzene to pro
cyclohexane is followed by air oxidation to yield a mixture of cyclohexanol
cyclohexanone (Figure 10). Nitric acid oxidation that yields adipic acid and nitrous ox
might make a measurable contribution to global warming and ozone depletion. An i
industry group was formed to facilitate development of new technologies for nitrous o
abatement.[56] So far all technologies have demonstrated the capability to very effici
abate nitrous oxide emissions.[56]

Inclusion of a *catA* gene encoding catechol 1,2-dioxygenase in the cate
producing microbe converts catechol into *cis,cis*-muconic acid (Figure 10).[57, 58] The
gene was isolated from *Acinetobacter calcoaceticus*.[59] *E. coli* WN1/pWN2.248
developed that synthesized 36.8 g/L of *cis,cis*-muconic acid in 22% (mol/mol) yield
glucose after 48 h of culturing under fed-batch fermentor conditions.[58] After fed-b

fermentation was complete, the cells were removed from the broth, which was treated with activated charcoal and subsequently filtered to remove soluble protein. Hydrogenation of the resulting solution with 10% Pt on carbon (5% mol/mol) at 3400 kPa of $H_2$ pressure for 2.5 h at ambient temperature afforded a 97% (mol/mol) conversion of *cis,cis*-muconic acid into adipic acid. [57, 58]

**Metabolic engineering of *E. coli* to increase product titers and yields**

To compete with chemical synthesis, microbial synthesis must produce value-added chemicals in a high yield (mol/mol percent conversion) and high titer (product concentration) while taking advantage of abundant, inexpensive feedstocks. Metabolic engineering of biocatalytic syntheses utilizing the shikimate pathway has focused on the elimination of feedback inhibition, overexpression of specific enzymes leading to the desired product, and elimination of enzymes and pathways resulting in depletion of the desired product. For example, in commercial tryptophan-producing strains, the *tna* gene encoding the tryptophan-degrading enzyme tryptophanase is deleted, and the genes in the terminal pathway are all overexpressed.[60]

Alteration of central metabolism in *E. coli* to channel increased carbon flow into the common pathway has been the focus of considerable research activity.[61] In addition to overexpressing feedback-insensitive DAHP synthase, strategies have been evaluated to increase E4P and PEP availability. The rate of aromatic amino acid biosynthesis is controlled by modulation of the catalytic activity of DAHP synthase. There are three isozymes of DAHP synthase expressed in *E. coli*, each of which is sensitive to feedback inhibition by one of the three aromatic amino acids. The genes *aroF*, *aroG* and *aroH* encode tyrosine-sensitive, phenylalanine-sensitive, and tryptophan-sensitive DAHP synthase isozymes, respectively. The in vivo activities of these enzymes are dictated by their expression levels, feedback inhibition by aromatic amino acids, and the availability of their E4P and PEP substrates.

Overexpression of DAHP synthase does not necessarily mean that the in viv activity of DAHP synthase has been increased because of the prominent regulatory rol played by feedback inhibition.[62] Several alleles that encode feedback insensitive DAH synthase have been obtained by mutation of the *aroF*,[63] *aroG*,[64] and *aroH*[65] loci. Extensiv mutation is not required to make the DAHP synthase feedback-insensitive. For instance, single amino acid change can render AroF catalytic activity insensitive to the concentratio of Tyr.[63a] Insensitivity to feedback inhibition by aromatic amino acids increases the in viv catalytic activity of each molecule of DAHP synthase.

In 1999, several different strategies were explored in this lab for circumventin transcriptional repression of *aroF*.[61f] One strategy placed plasmid-localized $aroF^{FBR}$ unde strong promoters such as $P_{tac}$. Another strategy utilized increased copies of $aroF^{FBR}$ an (or) its unmodified native promoter to titrate away the cellular supply of TyrR protein tha binds the promoter region of $aroF^{FBR}$ and represses transcription of $aroF^{FBR}$. This was th reason to design strategies to localize one $aroF^{FBR}$ locus per plasmid, two $aroF^{FBR}$ loci pe plasmid, and one $aroF^{FBR}$ locus accompanied by one $P_{aroF}$ promoter per plasmid.[6 However, higher DAHP synthase activity did not necessarily translate into higher shikimat pathway yields or titers as evidenced by the relationship between DAHP synthase specifi activity and DHS synthesized when plasmid-localized $aroF^{FBR}$ was under $P_{tac}$ control. [61f]

Ultimately, the catalytic activity of DAHP synthase increases to a point wher further amplification of even feedback-resistant DAHP synthase does not lead to improve synthesis of either aromatic amino acids or precursors to these end products. Historicall attention was focused on the in vivo availability of phosphoenolpyruvic acid (PEP) at th juncture.[61] A number of different cellular processes and enzymes compete with DAH synthase for PEP including pyruvate kinase, PEP carboxylase, and the PEP-depende carbohydrate:phosphotransferase system (PTS) for uptake of glucose and structurall related sugars. Efforts to improve intracellular PEP availability began with mutation inactivation of PEP carboxylase[61c, 66] and pyruvate kinase.[67] More recent efforts hav

focused on circumventing PTS-mediated uptake of glucose in microbes such as *E. coli* th
will be discussed in detail next. Discovery that E4P availability is a critical limiting factor
aromatic amino acid biosynthesis separated these two periods of research focusing c
intracellular PEP availability. [61a]

Mutational inactivation of PEP carboxylase and pyruvate kinase did not lead
improvements in aromatic amino acid biosynthesis that were significant or practical. Fo
example, mutational inactivation of *ppc* encoded PEP carboxylase results in a slow-growin
*E. coli ppc*$^-$ strain that requires succinic acid supplementation.[61c] Although *E. coli* ppc
produces 10-fold higher phenylalanine titers relative to *E. coli* with native expression c
*ppc*, these phenylalanine titers are 10-fold lower than the amount of acetic acid which
produced. Also, the 1.5 g/L titers of phenylalanine produced by *E. coli ppc*$^-$ are rathe
small relative to the 46 g/L of phenylalanine which can be produced by *E. coli*.[68]

In 1990, Frost and coworkers published the first work indicating that E4
availability was also an important factor limiting in vivo DAHP synthase activity.[69] As a
aldose phosphate which can not exist in solution in a cyclic form, E4P is prone t
dimerization, trimerization and polymerization.[70] Dissociation of these E4P forms back t
monomeric E4P is quite slow. This is probably the underlying reason why nature closel
matches the rate of E4P synthesis with the rate of E4P utilization, thereby maintaining lo
steady-state concentrations of E4P. Low steady-state concentrations of E4P likely limit th
substrate's availability thereby limiting in vivo activity of DAHP synthase. Inspection of th
enzymes which catalyze reactions where E4P is either a substrate or a product led to th
pentose phosphate pathway and the enzyme transketolase (Figure 11). Overexpression c
transketolase increases the levels of E4P available to the cell for channeling into th
common pathway.[69] Two of the three intracellular E4P-generating reactions are catalyze
by transketolase while the third reaction is catalyzed by transaldolase (Figure 11
Transketolase also serves to generate the substrate D-sedoheptulose-7-phosphate for th

E4P-generating reaction catalyzed by transaldolase. Thus, transketolase plays a major ro

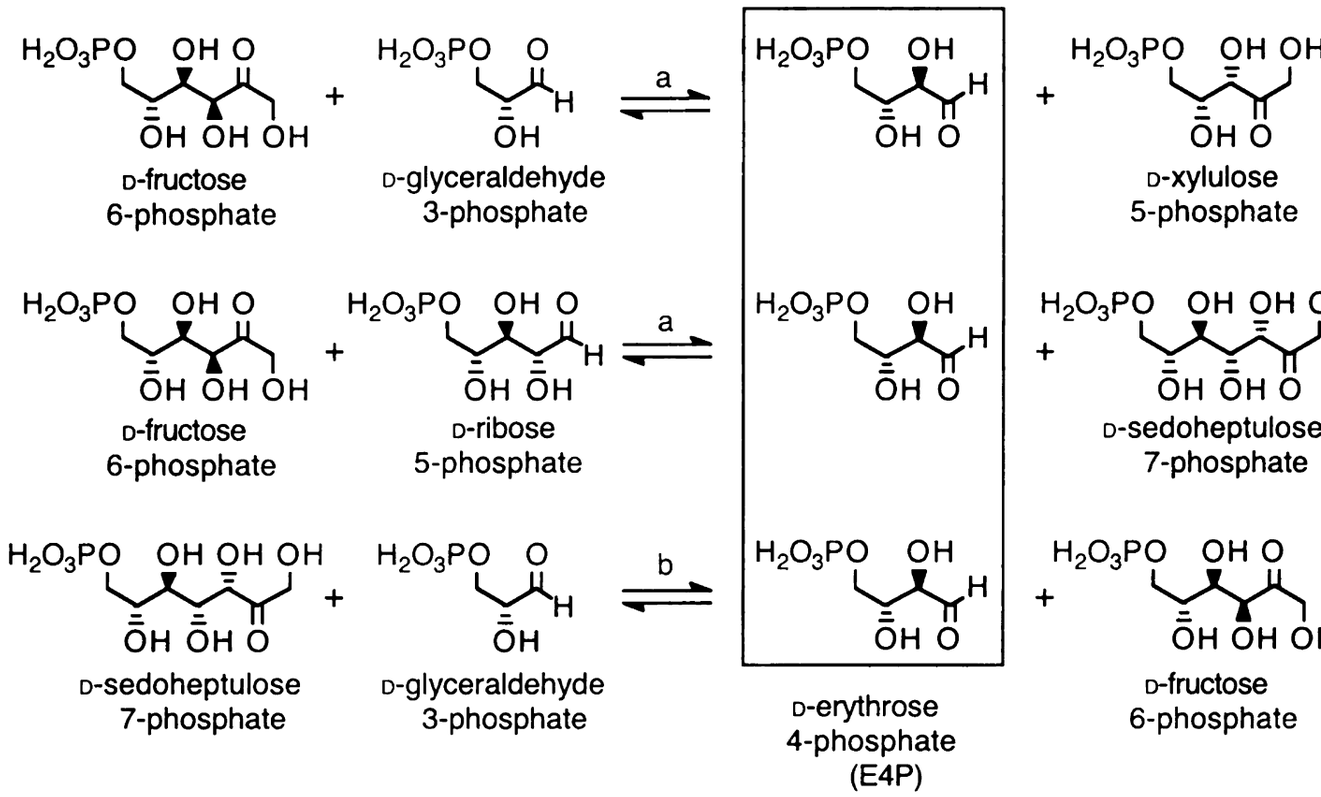



**Figure 11. Reactions catalyzed by (a) transketolase and (b) transaldolase.**

in increasing the levels of E4P available to the cell for aromatic production. DAH synthase, which catalyzes the first and irreversible reaction of the common pathwa commits elevated levels of carbon generated by transketolase to aromatic amino ac biosynthesis only if sufficient PEP is around. Work combining expression of feedbac insensitive *aroG* with overexpression of *tktA* has revealed that increased DAHP syntha catalytic activity in tandem with increased transketolase catalytic activity leads to a twofo increase in carbon flow directed into the common pathway above that achieved wi overexpression of DAHP synthase alone.[61f] By combining fed-batch fermentor contr with overexpression of a feedback-insensitive isozyme of DAHP synthase a overexpression of transketolase, shikimate pathway products were synthesized in 36 (mol/mol) yield from glucose,[61f] also a two-fold increase above the achieved yield wi overexpression of DAHP synthase alone. Overexpression of transaldolase also reliev E4P limitation in the presence of amplified PEP synthase, but no further improvements

aromatic amino acid biosynthesis are observed relative to when transketolase is also overexpressed.[71]

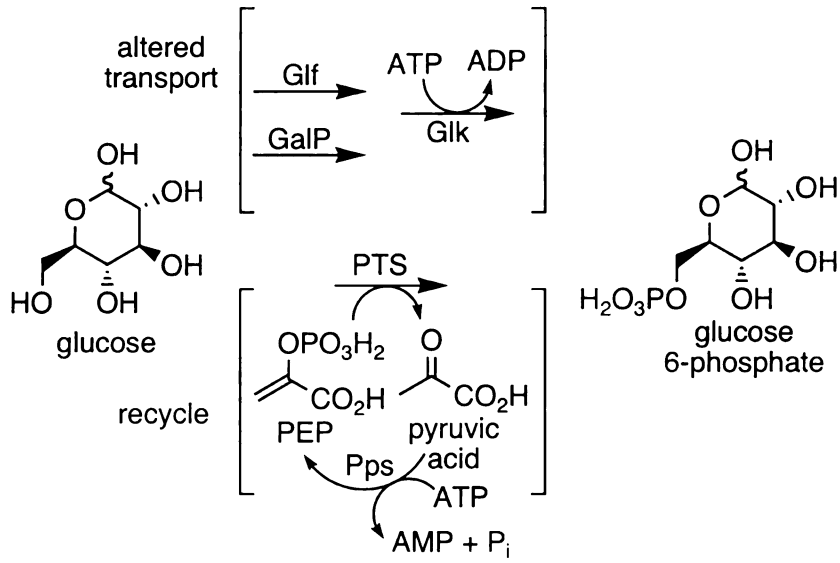



**Figure 12. Different strategies to remove PEP limitation.** Glucose facilitator (Glf), galactose permease (GalP), glucokinase (Glk), phosphoenolpyruvate synthase (Pps), phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS).

Liao and coworkers were the first to examine the impact of amplifying the expression of PEP synthase on shikimate pathway product yields.[61d] PEP synthase (Figure 12) catalyzes the generation of PEP from pyruvate with the expenditure of two of the high-energy, phosphodiester linkages of ATP to form AMP.[72] Using an *E. coli aroB* construct with overexpression of a feedback-insensitive isozyme of DAHP synthase, transketolase, and PEP synthase, DAH was synthesized in 90% (mol/mol) yield from glucose.[61d, 61e] This result indicated a strategy for circumventing the limitation imposed on product yields by the PTS employed by *E. coli* and many other microbes.

Some approaches to remove PEP limitation uses the principle of circumventing PTS-mediated glucose transport. One approach is to employ a different source of carbon for microbial synthesis of shikimate pathway products. For example, higher shikimate pathway product yields have been achieved when *E. coli* is cultured on D-xylose or L-

arabinose due to the transport of these pentoses in *E. coli* by permeases driven by the conversion of ATP to ADP.[61e, 73] Unfortunately, commercial D-xylose and L-arabinose stream that are pure, abundant, and inexpensive, are not currently available.

Another approach is to transport glucose by non-PTS mechanism including facilitated diffusion mediated by the *Z. mobilis glf*-encoded glucose facilitator and galactose-proton symport mediated by the *E. coli galP*-encoded galactose permease (Figure 12). Ingram and coworkers[74] were the first to demonstrate that *E. coli* mutants lacking PTS-mediated glucose transport and phosphorylation can grow on glucose with heterologous expression of the *Z. mobilis glf*-encoded glucose facilitator protein and *Z. mobilis glk*-encoded glucokinase (Figure 12). Heterologous expression of the *Z. mobilis glf*-encoded glucose facilitator[75] increased the concentrations of L-phenylalanine synthesized by various *E. coli* constructs.

Various groups studied the impact of replacement of PTS with the GalP galactose-proton symport system to transport glucose on shikimate pathway products yields in *E. coli* (Figure 12).[76] Valle and coworkers examined DAHP synthesis in *E. coli* NF9/pRW5tkt, which relied on GalP-mediated glucose transport with inactivation of PTS-mediated glucose transport.[76a] *E. coli* NF9/pRW5tkt synthesized 2.4-fold higher concentrations of DAHP under shake-flask culture conditions relative to *E. coli* PB103/pRW5tkt, which retained PTS-mediated glucose transport.[76a] Plasmid pRW5tkt contained *tktA* encoding transketolase and $aroG^{FBR}$ encoding feedback-insensitive DAHP synthase. A 71% (mol/mol) yield of DAHP and a detailed analysis of carbon metabolism have recently been reported[76d, 76e] for *E. coli* constructs where GalP was recruited for glucose transport. In another study, Bailey and coworkers[76c] examined phenylalanine synthesis in *E. coli* PPA316/pSY130-14, which also relied on GalP-mediated glucose transport with inactivation of PTS-mediated glucose transport. In contrast to Valle's result, *E. coli* PPA316/pSY130-14 under fermentor-controlled culture conditions synthesized only 0-67% of the phenylalanine synthesized by *E. coli* PPA305/pSY130-14, which relied on a PTS for

glucose transport.[76c] Plasmid pSY130-14 contained *aroF*$^{FBR}$ encoding feedback-insensitiv DAHP synthase and *pheA*$^{FBR}$ encoding chorismate mutase-prephenate dehydratase. Th different impact of GalP-mediated glucose transport on the synthesis of phenylalanine[76 relative to the synthesis of DAHP[76a] can be attributed to several factors. The Bailey study[76 did not overexpress either transketolase or transaldolase. Increased availability o phosphoenolpyruvate in *E. coli* is not reflected in increased shikimate pathway produc yields until the availability of D-erythrose 4-phosphate is increased with overexpression o transketolase.[61e] These studies[76a, 76c] also did not examine the microbial synthesis of th same shikimate pathway product, and microbes were cultivated under different conditions.

Another unique strategy to remove PEP limitation employs adjuncts that can b readily converted to PEP. For example, addition of succinic acid to glucose-limited culture of *E. coli* constructs overexpressing phosphoenolpyruvate carboxylase leads to synthesis o increased concentrations of shikimate pathway products.[77]

Most of the experiments mentioned above were performed in shake-flask and no under fed-batch fermentor conditions. Under typical shake-flask conditions, cells were firs grown to stationary phase in rich medium using antibiotics as the selection pressure t maintain plasmids. These cells were then harvested and resuspended in minimal medium Glucose in the minimal medium is then converted to the desired molecule with limited or n cell growth. Although this type of cell cultivation has been widely employed in evaluatin pathway engineering, there are several inherent problems associated with such experimenta procedures. For example, the use of rich medium and antibiotics can dramatically increas the cost if such culturing methods were scaled up. Centrifugation and resuspension of cell also complicate large-scale manufacture. Under shake-flask conditions, because cells ar grown in rich medium prior to resuspension in minimal medium, calculated yields do nc accurately reflect the percent conversion of carbohydrate into product. DAH yield exceeding the theoretical maximum yield were repeatedly reported when *aroG*$^{FBR}$, *tktA* an *pps* were overexpressed in an *E. coli aroB* cultured under shake-flask conditions.[61a, 61d, 61

In addition, under shake-flask conditions, cultures begin in glucose-rich environment and end in a glucose-deficient environment, and both oxygenation levels and pH are difficult to control.

In this thesis, fed-batch fermentor conditions as opposed to shake-flask culture conditions were employed to analyze the efficiency of conversion of glucose to product. Use of a fermentor allows glucose, dissolved oxygen concentrations, and pH to be easily controlled. One of the advantages of using fed-batch fermentor conditions is that it can increase cell density more than 10-fold relative to the shake-flask cultivation. In addition, fed-batch fermentor conditions overcome most of disadvantages of shake-flask conditions. For example, only minimal medium is used in the process, and the growth of the cells and the production of DHS occur simultaneously under fed-batch fermentor conditions. A strategy of using nutritional pressure to maintain the plasmid has also been developed. As a result, antibiotics are no longer used in the fermentor.

## Use of microarray technique to investigate the consequences of metabolic engineering

### Functional genomics for metabolic engineers

During the last decade of the 20th century, basic biological research underwent a major revolution as life scientists switched their level of analyses from studying the expression of single genes and proteins to studying large numbers of genes and gene products simultaneously. During the first portion of the genomics era, researchers concentrated on accumulating DNA sequence information from a variety of bacteria. As genomic information has become available on a broad range of organisms, a new postgenomics era has begun in which data is used as a resource base to characterize gene expression under different conditions. Because of the emphasis on gene function, this research is also referred to as functional genomics. The purpose of functional genomics is to use the information made available upon sequencing a genome to quantitatively determine

the spatial and temporal accumulation patterns of specific mRNAs, proteins, and importan metabolites using high-throughput technologies.[78]

For metabolic engineers who are largely interested in using living organisms t produce proteins and metabolites for commercial purposes, functional genomics provide new tools and approaches for understanding, modeling, and ultimately manipulatin organisms. Functional genomics relies heavily on three levels of high-throughput analyses transcriptomics (or RNA profiling) for measuring levels of mRNA, proteomics fo determining concentrations of individual proteins, and metabolomics (metabolite profiling for determining the amounts of important metabolites. Tradeoffs exist among the analytica power of these systems in the amounts of data generated and the usefulness of the result obtained. These tradeoffs can readily be visualized in the context of the central dogma o molecular biology, in which DNA can be transcribed into mRNAs, mRNAs translated int proteins, and proteins act to catalytically interconvert metabolites.

As you move through the central dogma from DNA to metabolites, the informatio becomes increasingly useful in terms of function. While DNA sequence indicates wha genes are present in a bacterium, mRNA measurements show which of theses genes ar expressed. Similarly, protein measurements identify those specific mRNAs that are bein translated and the amount of the specific enzymes that are present. This may or may not b reflected in the mRNA level. Finally the amount of metabolite present (particularly if flu information can be deduced) may be more important than determining the potential fo product formation as estimated by measuring enzyme levels. Unfortunately, while th information might become more valuable as we read through the central dogma, it become more difficult to obtain, often less quantitative, and certainly more fragmentary.

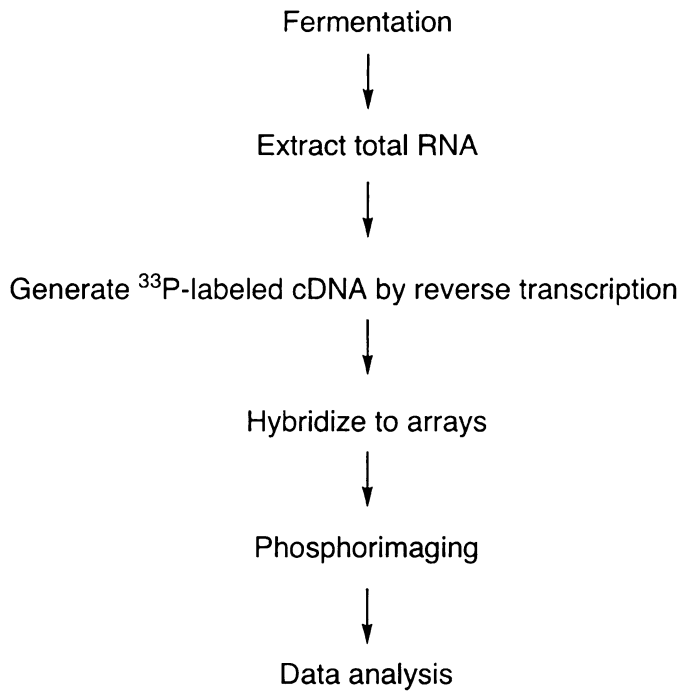



**Figure 13. Overview of Microarray process.**

**Microarray (RNA profiling)**

Analysis of the transcriptome in a given sample yields a measurement of the relative level of each mRNA. This level reflects the balance between the transcription rate of the gene and the rate at which that specific mRNA is being degraded. Several methods exist for determining the levels of cellular mRNAs; most utilize nucleic acid "probes" covalently bound to glass slides or nylon membranes. A major technology used is DNA microarray.[79] Microarrays are robotically printed sets of PCR products or conventionally synthesized oligonucleotides. High quality, pre-made oligonucleotide and PCR product arrays can be purchased from a number of sources, and increasing competitive pressure can be expected to continue to drive prices down and quality up. There are two major types of DNA microarrays. One is the oligonucleotide-based array and the other is the PCR product-based array. For example, the Panorama microbial gene arrays from Sigma Genosys make it possible for investigators to study expression at the transcriptional level of all open reading frames (ORFs) simultaneously in one experiment. The Panorama *E. coli*

gene array contains 4,290 PCR-amplified ORFs, representing all of the protein-encoding genes provided on a pair of duplicate nylon membranes. Affymetrix chips, on the other hand, synthesized high-density arrays of oligonucleotide probes (25-mers) *in situ* using light-directed chemistry.[80] The probes are designed to maximize sensitivity, specificity, and reproducibility, allowing consistent discrimination between specific and background signals, and between closely related target sequences. The major advantages of the Panorama *E. coli* gene array over Affymetrix chips are low capital investment and relatively easy operating procedure. Therefore, Panorama *E. coli* gene array was chosen as a platform to perform microarray analysis in our study. However, the Panorama *E. coli* gene array does not contain intergenic regions that include promoters or regions encoding small RNA molecules with possible regulatory functions.[81]

Thousands of gene probes can be represented on a single chip. mRNA is isolated from tissue or bacteria samples and used as a template to prepare a "target", for example, cDNA (or cRNA for Affymetrix chips) that is labeled, usually with fluorescent by active labeling compounds such as biotin, Cy3 and Cy5 or $P^{33}$ by incorporation radioactive nucleotides. The labeled target is hybridized to the probe on the microarray. Individual cDNAs/cRNAs from the target hybridize (bind) with the corresponding probe proportionally to their representation in a sample. A specialized scanner detects the hybridized molecules by fluorescence or radiolabel. The whole procedure used in this study is summarized in Figure 13.

**Application of microarray to metabolic engineering**

Traditional approaches of biocatalysis optimization use random screening, mutagenesis and engineering improvement. While these methods are still very effective, a better understanding of the underlying physiology using genomic tools can accelerate these efforts. Information obtained by the DNA microarrays can help pathway engineering and process optimization in several ways. (i) Regulatory circuitry and coordination of gene

expression among different pathways under different growth conditions can be measured by DNA microarray.[82] (ii) The physiological state of the cells during fermentation can be assessed by the genome-wide transcriptional patterns.[83] (iii) DNA microarray can help identify genes involved in a production process if they are coregulated.[84] (iv) The differences in genetic contents and expression profiles between wild-type and improved strains can be compared.[85] (v) The actual array data can be incorporated into the mathematical models to describe a cellular process.[86] Finally, general applications of DNA microarray technology to understand microbial physiology will continue to generate very large amounts of information that will ultimately benefit the pathway engineering and fermentation optimization effort.

Previous studies have reported differences in expression during the aerobic growth of *E. coli* on different carbon sources,[87] in minimal and complex media,[87] in high cell density culture conditions,[83] in response to heat shock,[88] in response to stress,[82b] and in response to pH change.[84] Additional studies have identified changes that contribute to ethanol tolerance in ethanologenic *Escherichia coli* by comparison of KO11 (parent) to LY01 (resistant mutant).[85] Expression levels for 205 of the ORFs were found to differ significantly between KO11 and LY01 under each of six different growth conditions. Statistical evaluation of differentially expressed genes according to various classification schemes identified physiological areas of importance. A large fraction of differentially expressed ORFs were globally regulated, leading to the discovery of a nonfunctional *fnr* gene in strain LY01. In agreement with a putative role for FNR in alcohol tolerance, increasing the copy number of *fnr* in KO11 decreased ethanol tolerance but had no effect on growth in the absence of ethanol. Other differences in gene expression provided additional clues that permitted experimentation. Tolerance appears to involve increased metabolism of glycine and increased production of betaine. Changes in the expression of several genes concerned with the synthesis of the cell envelope components were also noted, which may contribute to increased ethanol tolerance.

Expression arrays offer a unique opportunity to investigate the consequences of metabolic engineering and may provide clues to limitations in metabolic flux.[89] The fed-batch fermentation of *E. coli* under pH and oxygen controlled conditions is an attractive system for such studies due to the simplicity of culture conditions and the lack of complications from changes in oxygen availability or pH. In this thesis, the physiological basis for the increase in yield of shikimate pathway products[90] attendant with overexpression of PEP synthase was examined by gene array technology to detect differential transcription profiles of relative "high-yielding (51% mol shikimate pathway product/ mol glucose)" cells when PEP synthase was optimally overexpressed and relative "low-yielding (37% mol shikimate pathway product/ mol glucose)" cells when PEP synthase was not overexpressed. This work detailed the investigation of PEP synthase overexpression effect in *E. coli* under fed-batch fermentor conditions and the results demonstrated that the transcript abundance of a range of genes was altered by the single modification of phosphoenolpyruvate synthase gene expression. Several gene targets were identified for future strain development.

## Export Study

The question of how the desired product that is synthesized within the cell becomes excreted has been a neglected field until recently. Among other hypotheses, it was formerly assumed that unspecific leaks were present, or that diffusion enables the passage of the desired product into the culture medium. However, these rationales cannot explain why L-glutamate, for instance, is excreted "uphill" (i.e., at low intracellular concentrations toward the higher extracellular concentration).[91] It also cannot explain why a cellular discrimination between the physically very similar amino acids L-glutamate and L-aspartate occurs. Therefore, specific carriers must be present. It is now verified that at least the excretion of L-glutamate,[91] L-lysine,[92] L-isoleucine,[93] and L-threonine[94] by *C. glutamicum*, of cysteine[95] and L-threonine[96] by *E. coli,* are carrier-mediated processes.

A special situation prevails for the L-isoleucine export of *C. glutamicum*.[93] Because this amino acid is hydrophobic, both active export and diffusion actually contribute to the transport. This set of properties has serious consequences for the removal of L-isoleucine from the cell. Since diffusion is not a vectorial process, a high diffusion flux via the membrane always takes place at high L-isoleucine concentrations, irrespective of whether from the inside or outside of the cell. Therefore, in addition to exporting the L-isoleucine synthesized, the carrier activity must also be sufficient to export the L-isoleucine that reentered the cell by diffusion. Otherwise the exposure of the interior of the cell to high concentrations of the hydrophobic amino acid might influence vital functions. It is remarkable in this context that for the more hydrophobic aromatic and branched-chain amino acids, far lower amino acid concentrations are obtained in productions than with the charged amino acids, where extremely high concentrations (>150 g/L) are reached.

Shikimic acid synthesis affords an opportunity to elaborate an efflux system that is likely uniquely different from previously elaborated efflux systems given the structurally distinct class of molecule exported. Several bacterial proteins responsible for efflux of specific amino acids have recently been identified.[97] The LysE protein responsible for L-lysine export in *Corynebacterium glutamicum* is perhaps the best characterized amino acid efflux system.[92] LysE is a 25 kDa protein, which appears to span the cellular membrane six times. *C. glutamicum lysE* mutants fail to accumulate L-lysine extracellularly and have high intracellular lysine concentrations. Conversly, *C. glutamicum* strains possessing multiple copies of *lysE* export L-lysine at a rate five-fold higher than wild-type *C. glutamicum* and have relatively low intracellular lysine concentrations. Efflux of cystein and several related metabolites is encoded by *orf299* in *E. coli,*[95] and the *rgtB* locus in this same microbe has been found to participate in efflux of homoserine.[96]

Shikimic acid is a hydrophilic molecule and is likely transported out of the cell by a carrier-mediated process instead of diffusion. The identification of the carrier gene and

study of its regulation and mechanism is central to our goal of achieving high titer, high-yield syntheses of shikimic acid.

## REFERENCE


1 Industry of the Future: New Biocatalysis: Essential tools for a Sustainable 21st Century Chemical Industry. 1999, www.oit.doe.gov/chemicals/pdfs/ biocatalysis_roadmap.pdf.

2 PRODUCTION: DOWN BUT NOT OUT. *Chem. Eng. News* **2001**, *80*, 60-65.

3 Lewis, R. J. *Carcinogenically Active Chemicals*, Van Nostrand Reinhold, New York, 1991, p. 68.

4 Ember, L. Toxic Air Emissions - Tough EPA Rule Will Hit 112 Chemicals. *Chem. Eng. News* **1994**, *72*, 4.

5 (a) Yoshida, J.; Inomata, M. Trends in developments of aromatics production technologies. *Aromatikkusu* **2002**, *54*, 123-135. (b) Tullo, A. H. A New Source. *Chem. Eng. News* **2003**, *81*, 16-17.

6 (a) Frost, J. W.; Lievense, J. Prospects for Biocatalytic Synthesis of Aromatics in the 21st- Century. *New J. Chem.* **1994**, *18*, 341. (b) Bongaerts, J.; Kramer, M.; Muller, U.; Raeven, L.; Wubbolts, M. Metabolic Engineering for Microbial Production of Aromatic Amino Acids and Derived Compounds. *Metabol. Eng.* **2001**, *3*, 289-300.

7 (a) Haslam, E. *The Shikimic Pathway*; Wiley: New York, 1974. (b) Bentley, R. The Shikimate Pathway- a Metabolic Tree with Many Branches. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 307-384. (c) Herrmann, K. M. In *Amino Acids: Biosynthesis and Genetic Regulation*; Herrmann, K. M., Somerville, R. L., Ed.; Addison-Wesley: Reading, 1983: p. 301.

8 Frost, J. W.; Bender, J. L.; Kadonaga, J. T.; Knowles, J. R. Dehydroquinate synthase from *Escherichia coli*: purification, cloning, and construction of overproducers of the enzyme. *Biochemistry* **1984**, *23*, 4470-4475.

9 Carpenter, E. P.; Hawkins, A. R.; Frost, J. W.; Brown, K. A. Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis. *Nature* **1998**, *394*, 299-302.

10 (a) Smith, B. W.; Turner, M. J.; Haslem, E. Shikimate Pathway .4. Stereochemistry of 3-Dehydroquinate Dehydratase Reaction and Observations on 3-Dehydroquinate Synthetase. *J. Chem. Soc., Perkins Trans. 1* **1975**, *1*, 52-55. (b) Duncan, K.; Chaudhuri, S.; Campbell, M. S.; Coggins, J. R. The Overexpression and Complete Amino-Acid-Sequence of *Escherichia coli* 3-Dehydroquinase. *Biochem. J.* **1986**, *238*, 475-483.

11 (a) Anton, I. A.; Cogins, J. R. Sequencing and Overexpression of the *Escherichia coli* AroE Gene Encoding Shikimate Dehydrogenase. *Biochem. J.* **1988,** *249*, 319-326. (b) Chaudhuri, S.; Coggins, J. R. The Purification of Shikimate Dehydrogenase from *Escherichia coli. Biochem. J.* **1985**, *226*, 217-223.

12 (a) DeFeyter, R. C.; Pittard, J. Genetic and Molecular Analysis of AroL, the Gene for Shikimate Kinase-II in *Escherichia coli* K-12. *J. Bacteriol.* **1986**, *165*, 226-232. (b) DeFeyter, R. C.; Pittard, J. Purification and Properties of Shikimate Kinase-II from *Escherichia coli* K-12. *J. Bacteriol.* **1986**, *165*, 331-333.

13 LØbner-Olesen, A.; Marinus, M. G. Identification of the Gene (Arok) Encoding Shikimic Acid Kinase-I of *Escherichia coli*. *J. Bacteriol.* **1992**, *174*, 525-529.

14 Duncan, K.; Coggins, J. R. The Serc-AroA Operon of *Escherichia coli*: a Mixed Function Operon Encoding Enzymes from 2 Different Amino-Acid Biosynthetic Pathways. *Biochem. J.* **1986**, *234*, 49-57. (b) Duncan, K.; Lewendon, A.; Coggins. J. R. The Purification of 5-Enolpyruvylshikimate 3-Phosphate Synthase from an Overproducing Strain of *Escherichia coli*. *FEBS Lett.* **1984**, *165*, 121-127.

15 White, P. J.; Millar, G.; Coggins, J. The Overexpression, Purification and Complete Amino-Acid Sequence of Chorismate Synthase from *Escherichia coli*-K12 and Its Comparison with the Enzyme from *Neurospora crassa*. *Biochem. J.* **1988**, *251*, 313-322.

16 Hodgson J. BULK Bulk Amino-Acid Fermentation - Technology and Commodity Trading. *Biotechnology* **1994**, *12*, 152-155.

17 Chemical Economics Handbook, SRI International, 1999.

18 (a) Murdock, D.; Ensley, B. D.; Serdar, C.; Thalen, M. Construction of Metabolic Operons Catalyzing the Denovo Biosynthesis of Indigo in *Escherichia coli*. *Biotechnology* **1993**, *11*, 381-386. (b)Berry, A.; Battist, S.; Chotani, G.; Dodge, T. C.; Peck, S.; Power, S.; Weyler, W. Biosynthesis of indigo using recombinant *E. coli*: development of a biological system for the cost-effective production of a large volume chemical. Proceedings-Biomass Conference Americas: Energy, Environment, Agriculture and Industry, 2nd, Portland, Oreg. Aut. 21-24, 1995, p 1121-1129. (c) Berry, A.; Dodge, T. C.; Pepsin, M.; Weyler, W. Application of metabolic engineering to improve both the production and use of biotech indigo. *Indust. Microbiol. Biotech.* **2002**, *28*, 127-133.

19 Haq, I. U.; Ali, S. Microbiological Transformation of L-Tyrosine to 3,4-Dihydroxyphenyl L-Alanine (L-Dopa) by a Mutant Strain of *Aspergillus*. **2002**, *45*, 88-93.

20 De Boer, L.; Dijkhuizen, L. Microbial and enzymatic processes for L-phenylalanine production. *Adv. Biochem. Eng./Biotechnol.* **1990**, *41*, 1-27.

21 Takagi, M.; Nishio, Y.; Oh, G.; Yoshida, T. Control of L-phenylalanine production by dual feeding of glucose and L-tyrosine. *Biotechnol. Bioeng.* **1996**, *52*, 653-660.

22 Maiti, T. K.; Roy, A.; Mukherjee, S. K. Chatterjee, S. P. Microbial production of L-Tyrosine: A review. *Hindustan Antibiot. Bull.* **1995**, *37*, 51-65.

23 (a) Ito, H.; Sato, K.; Enei, H.; Hirose, Y. Improvement of microbial production of L-Tyrosine by gene dosage effect of *aroL* gene encoding shikimate kinase. *Agric. Biol. Chem.* **1990**, *54*, 823-824. (b) Ikeda, M.; Katsumate, R. Metabolic engineering to produce tyrosine or phenylalanine in a tryptophan-producing *Corynebacterium glutamicum* strain. *Appl. Environ. Microbiol.* **1992**, *58*, 781-785. (c) Ikeda, M.; Okamoto, K.; Katsumata, R. Cloning of the transketolase gene and the effect of its dosage on aromatic amino acid production in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* **1999**, *51*, 201-206.

24 Berry, A. Improving production of aromatic compounds in *Escherichia coli* by metabolic engineering. *Trends Biotechnol.* **1996**, *14*, 250-256.

25 Katsumata, R.; Ikeda, M. Hyperproduction of tryptophan in *Corynebacterium glutamicum* by pathway engineering. *Biotechnology* **1993**, *11*, 921-925.

26 Kurahashi, O.; Tsuchida, T.; Kawashima, N.; Ei, H.; Yamane, K. L-Tryptophan production by transformed *Bacillus subtilis*. 1984, JP 61096990.

27 (a) Tan, D. S.; Foley, M. A.; Shair, M. D.; Schreiber, S. L. Stereoselective Synthesis of over Two Million Compounds Having Structural Features Both Reminiscent of Natural Products and Compatible with Miniaturized Cell-Based Assays. *J. Am. Chem. Soc.* **1998**, *120*, 8565-8566. (b) Tan, D. S.; Foley, M. A.; Stockwell, B. R.; Shair, M. D.; Schreiber, S. L. Synthesis and Preliminary Evaluation of a Library of Polycyclic Small Molecules for Use in Chemical Genetic Assays. *J. Am. Chem. Soc.* **1999**, *121*, 90073-9087.

28 Kim, C. U.; Lew, W.; Williams, M. A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Mendel, D. B.; Tai, C. Y.; Laver, W. G.; Stevens, R. C. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: Design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.* **1997**, *119*, 681-690.

29 Weber, W. F. Hoffmann-La Roche, Ltd., personal communication.

30 Leuenberger, H. Hoffmann-La Roche, Ltd., personal communication.

31 Knop, D. R.; Draths, K. M.; Chandran, S. S.; Barker, J. L.; von Daeniken, R.; Weber, W.; Frost, J. W. Hydroaromatic Equilibration During Biosynthesis of Shikimic Acid. *J. Am. Chem. Soc.* **2001**, *123*, 10173-10182.

32 Draths, K. M.; Knop, D. R.; Frost, J. W. Shikimic Acid and Quinic Acid: Replacing Isolation from Plant Sources with Recombinant Microbial Biocatalysis. *J. Am. Chem. Soc.* **1999**, *121*, 1603-1604.

33 Kirsch, M. A.; Williams, D. J. Understanding the Thermoplastic Polyester Business. *CHEMTECH* **1994**, *24*, 40-49.

34 Szmant, H. H. In *Organic Building Blocks of the Chemical Industry*; New York: Wiley, 1989, p. 467.

35 Lenga, R. E.; Votoupal, K. L. *The Sigma-Aldrich Libray of Regulatory and Safety Data*; Sigma-Aldrich: Milwaukee, WI, 1993.

36 Siebert, M.; Berchthold, A.; Melzer, M.; May, U. Berger, U. *FEBS Lett.* Ubiquinone Biosynthesis - Cloning of the Genes-Coding for Chorismate Pyruvate-Lyase and 4-Hydroxybenzoate Octaprenyl Transferase from *Escherichia coli.* **1992**, *307*, 347-350.

37 Barker, J.; Draths, K. D.; Frost, J. W. Microbial synthesis of p-hydroxybenzoic acid from glucose. *Biotechnol. Bioeng.* **2001**, *76*, 376-390.

38 Draths, K. M.; Ward, T. L.; Frost, J. W. Biocatalysis and 19th-Century Organic-Chemistry - Conversion of D-Glucose into Quinoid Organics. *J. Am. Chem. Soc.* **1992**, *114*, 9725-9726.

39 (a) Hanessian, S.; Beaulieu, P.; Dube, D. A Novel Synthetic Route to the Hexahydrobenzofuran Subunit of the Avermectins and Milbemycins. *Tetrahedron lett.*

**1986**, *27*, 5071-5074. (b) Flack, J. R.; Yadagiri, P. Enantiospecific Synthesis of D-Myo-Inositol 1,4,5-Trisphosphate from (-)-Quinic Acid. *J. Org. Chem.* **1989**, *54*, 5851-5852. (c) Molin, H. ; Pring, B. G. Regioselective and Stereoselective Synthesis of the Carbocyclic Analog of 3-Deoxy-Beta-D-*manno*-2-Octulopyranosonic Acid (Beta- Kdo) from (-)-Quinic Acid. *Tetrahedron lett.* **1985**, *26*, 677-680. (d) Montchamp, J.-L.; Piehler, L. T.; Frost, J. W. Diastereoselection and In vivo Inhibition of 3-Dehydroquinate Synthase. *J. Am. Chem. Soc.* **1992**, *114*, 4453-4459. (e) Widlanski, T.; Bender, S. L.; Knowles, J. R. Dehydroquinate Synthase: a Sheep in Wolfs Clothing. *J. Am. Chem. Soc.* **1989**, *111*, 2299-2300.

40 Ran, N.; Knop, D. R.; Draths, K. M.; Frost, J. W. Benzene-Free synthesis of Hydroquinone. *J. Am. Chem. Soc.* **2001**, *123*, 10927-10934.

41 Szmant, H. H. In *Organic Building Blocks of the Chemical Industry*; New York: Wiley, 1989, p. 512.

42 Krumenacker, L.; Constantini, M.; Pontal, P.; Sentrnac, J. In *Kirt-Othmer encyclopedia of chemical technology*, 4th ed. Kroschwitz, J. I., Howe-Grant, M., Eds.; Wiley, New York, 1995, vol 13, p 996.

43 Chang, Y.; Almy, E. A.; Blamer, G. A.; Gray, J. I.; Frost, J. W.; Strasburg, G. M. Antioxidant Activity of 3-Dehydroshikimic Acid in Liposomes, Emulsions, and Bulk Oil. *J. Agric. Food Chem.* **2003**, *51*, 2753-2757.

44 Knop, D. R.; Draths, K. M.; Chandran, S. S.; Barker, J. L.; von Daeniken, R.; Weber, W.; Frost, J. W. Hydroaromatic Equilibration During Biosynthesis of Shikimic Acid. *J. Am. Chem. Soc.* **2001**, *123*, 10173-10182.

45 Kambourakis, S.; Draths, K. M.; Frost, J. W. Synthesis of Gallic Acid and Pyrogallol from Glucose: Replacing Natural Product Isolation with Microbial Catalysis. *J. Am. Chem. Soc.* **2000**, *122*, 9042-9043.

46 Szmant, H. H. In *Organic Building Blocks of the Chemical Industry*; New York: Wiley, 1989, p. 519.

47 Krumenacker, L.; Costantini, M.; Pontal, P.; Sentenac, J. In *Kirt-Othmer encyclopedia of chemical technology*, Kroschwitz, J. I., Howe-Grant, M., Eds.; Wiley, New York, 1995, vol 13, p 996.

48 (a) Franck, H.-G.; Stadelhofer, J. W. Industrial Aromatic Chemistry; Springer-Verlag: New York, 1988; p. 183-190. (b) Varagnat, J. In *Kirt-Othmer encyclopedia of chemical technology*, 3rd ed.; Grayson, M., Ed.; Wiley: New York, 1981; vol. 13, p. 39-69. (c) Szmant, H. H. In *Organic Building Blocks of the Chemical Industry*; New York: Wiley, 1989, p. 512-519.

49 Draths, K. M.; Frost, J. W. Environmentally Compatible Synthesis of Catechol from D-Glucose. *J. Am. Chem. Soc.* **1995**, *117*, 2395-2400.

50 Walton, N. J.; Mayer, M. J.; Narbad, A. Molecular of Interest: Vanillin. *Phytochemistry* **2003**, *63*, 505-515.

51 Makkar, H. P. S.; Beeker, K. Isolation of tannins from leaves of some trees and shrubs and their properties. *J. Agric. Food Chem.* **1994**, *42*, 731-734.

52 (a) Esposito, L.; Formanek, K.; Kientz, G.; Mauger, F.; Maureaux, V.; Robert, G.; Truchet, F. In *Kirk-Othmer Encyclopedia of Chemical Technology*; Fourth Ed., Kroschwitz, J. I.; Howe-Grant, M., Ed.; Wiley: New York, 1997; Vol. 24, p 812. (b) Van Ness, J. H. In *Kirk-Othmer Encyclopedia of Chemical Technology,* 3rd Edition; Wiley: New York, 1983, Volum 23, p. 704.

53 Li, K.; Frost, J.W. Synthesis of Vanillin from Glucose. *J. Am. Chem. Soc.* **1998**, *120*, 10545-10546.

54 World Nylon 6 & 66 Supply/Demand Report. http://www.thepcigroup.com/fibres/

55 Thiemens, M. H.; Trogler, W. C. Nylon Production - an Unknown Source of Atmospheric Nitrous- Oxide. *Science* **1991**, *251*, 932-934.

56 Reimer, R. A.; Slaten, C. S.; Seapan, M.; Koch, T. A.; Triner, V. G. Adipic acid industry- $N_2O$ abatement implementation of technologies for abatement of $N_2O$ emission associated with adipic acid manufacture. Non-$CO_2$ Greenhouse Gases: Scientific Understanding, Control and Implementation, Proceedings of the International Symposium, 2nd, Noordwijkerhout, Netherlands, sept. 8-10, 1999, 347-358.

57 Draths, K. M.; Frost, J. W. Environmentally Compatible Synthesis of Adipic Acid from D-glucose. *J. Am. Chem. Soc.* **1994**, *116*, 399-400.

58 Niu, W.; Draths, K. M.; Frost, J. W. Benzene-Free Synthesis of Adipic Acid. *Biotechnol. Prog.* **2002**, *18,* 201-211.

59 Neidle, E. L.; Ornston, L. N. Cloning and Expression of *Acinetobacter calcoaceticus* Catechol 1,2-Dioxygenase Structural Gene Cata in *Escherichia coli. J. Bacteriol.* **1986**, *168*, 815-820.

60 (a) Aiba, S.; Tsunekawa, H.; Imanaka, T. New Approach to Tryptophan Production by *Escherichia coli*: Genetic Manipulation of Composite Plasmids Invitro. *Appl. Env. Microbiol.* **1982**, *43*, 289-297. (b) Tsunekawa, H.; Imanaka, T.; Aiba, S. Phenotypic Stability of Trp Operon Recombinant Plasmids in *Escherichia coli. J. Gen. Microbiol.* **1980**, *118*, 253-261.

61 (a) Draths, K. M.; Pompliano, D. L.; Conley, D. L.; Frost, J. W.; Berry, A.; Disbrow, G. L.; Staversky, R. J.; Lievense, J. C. Biocatalytic Synthesis of Aromatics from D-Glucose: The Role of Transketolase. *J. Am. Chem. Soc.* **1992**, *114*, 3956-3962. (b) Gubler, M., Jetten, M., Lee, S. H., Sinskey, A. J. Cloning of the Pyruvate-Kinase Gene (Pyk) of *Corynebacterium glutamicum* and Site-Specific Inactivation of Pyk in a Lysine-Producing *Corynebacterium lactofermentum* Strain. *Appl. Env. Microbiol.* **1994,** *60*, 2494-2500. (c) Miller, J. E.; Backman, K.C.; O'Conner, M. J.; Hatch, R. T. Production of Phenylalanine and Organic-Acids by Phosphoenolpyruvate Carboxylase-Deficient Mutants of *Escherichia coli. J. Ind. Microbiol.* **1987,** 2, 143-149. (d) Patnaik, R.; Liao, J. C. Engineering of *Escherichia coli* Central Metabolism for Aromatic Metabolite Production with Near Theoretical Yield. *Appl. Environ. Microbiol.* **1994,** *60*, 3903-3908. (e) Patnaik, R.; Spitzer, R. G.; Liao, J. C. Pathway Engineering for Production of Aromatics in *Escherichia coli*: Confirmation of Stoichiometric Analysis by Independent Modulation of AroG, TktA, and Pps Activities. *Biotechnol. Bioeng.* **1995**, *46*, 361-370. (f) Li, K.; Mikola, M. R.; Draths, K. M.; Worden, R. M.; Frost, J. W. Fed-Batch Fermentor Synthesis of 3-Dehydroshikimic Acid Using Recombinant *Escherichia coli. Biotechnol. Bioeng.* **1999**, *64*,

61-73.

62 Ogino, T.; Garner, C.; Markley, J. L. ; Herrmann, K. M. Biosynthesis of Aromatic-Compounds - $C^{13}$ NMR-Spectroscopy of Whole *Escherichia coli* Cells. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5828-5832.

63 (a) Weaver, L. M.; Herrmann, K. M. Cloning of an AroF Allele Encoding a Tyrosine-Insensitive 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase. *J. Bacteriol.* **1990**, *172*, 6581. (b) Mikola, M. R.; Widman, M. T.; Worden, R. M. In situ mutagenesis and chemotactic selection of microorganisms in a diffusion gradient chamber. *Appl. Biochem. Biotechnol.* **1998**, *70-72*, 905-918.

64 (a) Kikuchi, T.; Sotochi, N.; Fukase, K.; Kojima, H.; Kurahashi, O.; matsui, Y. Aromatic amino acid manufacture enhancement with recombinant microorganism. JP Patent 04248983, 1993. (b) Tonouchi, N.; Kojima, H.; Matsui, H. The use of feedback-insensitive enzymes in the production of aromatic amino acids by fermentation. Eur. Pat. Appl. 488424, 1992.

65 Ray, J. M.; Yanofsky, C.; Bauerle, R. Mutational Analysis of the Catalytic and Feedback Sites of the Tryptophan-Sensitive 3-Deoxy-D-*Arabino*-Heptulosonate-7-Phosphate Synthase of *Escherichia coli. J. Bacteriol.* **1988**, *170*, 5500-5506.

66 Backman, K.C. U.S. Patent 5,169,768, 1992.

67 Mori, M.; Yokota, A. ; Sugitomo, S.; Kawamura, K. Patent JP 62,205,782, 1987.

68 (a) Konstantinov, K. B.; Nishio, N.; Yoshida, T. Glucose Feeding Strategy Accounting for the Decreasing Oxidative Capacity of Recombinant *Escherichia coli* in Fed-Batch Cultivation for Phenylalanine Production. *Ferment. Bioeng.* **1990**, *70*, 253-260. (b) Konstantinov, K. B.; Nishio, N.; Seki, T.; Yoshida, T. Physiologically Motivated Strategies for Control of the Fed-Batch Cultivation of Recombinant *Escherichia coli* for Phenylalanine Production. *Ferment. Bioeng.* **1991**, *71*, 350-355.

69 (a) Draths, K. M.; Frost, J. W. Synthesis Using Plasmid-Based Biocatalysis: Plasmid Assembly and 3-Deoxy-D-*Arabino*-Heptulosonate Production. *J. Am. Chem. Soc.* **1990**, *112*, 1657-1659. (b) Frost, J. W. U.S. Patent 5,168,056, 1992.

70 Williams, J. F.; Blackmore, P. F.; Duke, C. C. MacLeod, J. K. fact, uncertainty and speculation concerning the biochemistry of D-erythrose-4-phosphate and its metabolic roles. *Int. J. Biochem.* **1980**, *12*, 339-344.

71 (a) Farabaugh, M. M.S. Thesis, Michigan State University, April 1996. (b) Lu, J.-L.; Liao, J. C. Metabolic engineering and control analysis for production of aromatics: Role of transaldolase. *Biotechnol. Bioeng.* **1997**, *53*, 132-138.

72 Niersbach, M.; Kreuzaler, F.; Geerse, R. H.; Postma, P. W.; Hirsch, H. J. Cloning and nucleotide sequence of the *Escherichia coli* K-12 *ppsA* gene, encoding PEP synthase. *Mol. Gen. Genetics* **1992**, *231*, 332-336.

73 Li, K.; Frost, J. W. Microbial Synthesis of 3-Dehydroshikimic Acid: A Comparative Analysis of D-Xylose, L-Arabinose, and D-Glucose Carbon Sources. *Biotechnol. Prog.* **1999**, *15*, 876-883.

74 (a) Parker, C.; Barnell, W. O.; Snoep, J. L.; Ingram, L. O.; Conway, T. Characterization of the *Zymomonas mobilis* glucose facilitator gene product (*glf*) in recombinant *Escherichia coli*: Examination of transport mechanism, kinetics and the role of glucokinase in glucose transport. *Mol. Microbiol.* **1995**, *15*, 795-802. (b) Snoep, J. L.; Arfman, N.; Yomano, L. P.; Fliege, R. K.; Conway, T.; Ingram, L. O. Reconstruction of Glucose Uptake and Phosphorylation in a Glucose-Negative Mutant of *Escherichia coli* by Using *Zymomonas mobilis* Genes Encoding the Glucose Facilitator Protein and Glucokinase. *J. Bacteriol.* **1994**, *176*, 2133-2135. (c) Weisser, P.; Krämer, R.; Sahm, H.; Sprenger, G. A. Functional Expression of the Glucose Transporter of *Zymomonas mobilis* Leads to Restoration of Glucose and Fructose Uptake in *Escherichia coli* Mutants and Provides Evidence for Its Facilitator Action. *J. Bacteriol.* **1995**, *177*, 3551-3554.

75 (a) Sprenger, G.; Sahm, H.; Karutz, M.; Sonke, T. Microbial Preparation of Substances from Aromatic Metabolism/II. WO 98/18937, 1998. (b) Kraemer, M.; Karutz, M.; Sprenger, G.; Sahm, H. Microbial Preparation of Substances from Aromatic Metabolism/III. WO 99/55877, 1999. (c) Sprenger, G.; Siewe, R.; Sahm, H.; Karutz, M.; Sonke, T. Microbial Preparation of Substances from Aromatic Metabolism/I. U.S. Patent 6,316,232 B1, 2001.

76 (a) Flores, N.; Xiao, J.; Berry, A.; Bolivar, F.; Valle, F. Pathway Engineering for the Production of Aromatic Compounds in *Escherichia coli*. *Nature Biotechnol.* **1996**, *14*, 620-623. (b) Chen, R.; Yap, W. M. G. J.; Postma, P. W.; Bailey, J. E. Comparative Studies of *Escherichia coli* Strains using Different Glucose Uptake Systems: Metabolism and Energetics. *Biotechnol. Bioeng.* **1997**, *56*, 583-590. (c) Chen, R.; Hatzimanikatis, V.; Yap, W. M. G. J.; Postma, P. W.; Bailey, J. E. Metabolic Consequences of Phosphotransferase (PTS) Mutation in a Phenylalanine-Producing Recombinant *Escherichia coli*. *Biotechnol. Prog.* **1997**, *13*, 768-775. (d) Baez, J. L.; Bolivar, F.; Gosset, G. Determination of 3-Deoxy-D-*Arabino*-Heptulosonate 7-Phosphate Productivity and Yield from Glucose in *Escherichia coli* Devoid of the Glucose Phosphotransferase Transport System. *Biotechnol. Bioeng.* **2001**, *73*, 530-535. (e) Flores, S.; Gosset, G.; Flores, N.; de Graaf, A. A.; Bolivar, F. Analysis of Carbon Metabolism in *Escherichia coli* Strains with an Inactive Phosphotransferase System by $^{13}C$ Labeling and NMR Spectroscopy. *Metabol. Eng.* **2002**, *4*, 124-137.

77 Li, K.; Frost, J. W. Utilizing Succinic Acid as a Glucose Adjunct in Fed-Batch Fermentation: Is Butane a Feedstock Option in Microbe-Catalyzed Synthesis? *J. Am. Chem. Soc.* **1999**, *121*, 9461-9462.

78 Oliver, D. J.; Nikolau, B.; Wurtele, E. S. Functional Genomics: High-Throughput mRNA, Protein, and Metabolite Analyses. *Met. Engin.* **2002**, *4*, 98-106.

79 Brown, P. O.; Botstein, D. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **1999**, *21*, 33-37.

80 (a) Eisen, M. B.; Brown, P. O. DNA arrays for analysis of gene expression. *Methods Enzymol.* **1999**, *303*, 179-205. (b) Lipshutz, R. J.; Fodor, S. P.; Gingeras, T. R.; Lockhart, D. J. High density synthetic oligonucleotide arrays. *Nat. Genet.* **1999**, *21*, 20-24.

81 Lease, R. A.; Belfort, M. A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 9919-9924.

82 (a) Zheng, M.; Wang, X.; Templeton, L. J.; Smulski, D. R.; LaRossa, R. A.; Storz, G. DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J. Bacteriol.* **2001**, *183*, 4562-4570. (b) Tucker, D. L.; Tucker, N.; Conway, T. Gene expression profiling of the pH response in *Escherichia coli. J. Bacteriol.* **2002**, *184*, 6551-6558.

83 Yoon, S. H.; Han, M.-J.; Lee, S. Y.; Jeong, K. J.; Yoo, J.-S. Combined transcriptome and proteome analysis of *Escherichia coli* during high cell density culture. *Biotechnol. Bioeng.* **2003**, *81*, 753-767.

84 Gonzalez, R.; Tao, H.; Shanmugam, K. T.; York, S. W.; Ingram, L. O. Global Gene Expression Differences Associated with Changes in Glycolytic Flux and Growth Rate in *Escherichia coli* During the Fermentation of Glucose and Xylose. *Biotechnology Progress* **2002**, *18*, 6-20.

85 Gonzalez, R.; Tao, H.; Purvis, J. E.; York, S. W.; Shanmugam, K. T.; Ingram, L. O. Gene Array-Based Identification of Changes That Contribute to Ethanol Tolerance in Ethanologenic *Escherichia coli*: Comparison of Ko11 (Parent) to Ly01 (Resistant Mutant). *Biotechnol. Prog.* **2003**, *19*, 612-623.

86 Sabatti, C.; Rohlin, L.; Oh, M.-K.; Liao, J. C. Co-expression pattern from DNA microarray experiments as a tool for operon prediction. *Nucl. Acid Res.* **2002**, *30*, 2886-2893.

87 (a) Tao, H.; Bausch, C.; Richmond, C.; Blattner, F. R.; Conway, T. Functional genomics: expression analysis of *Escherichia coli* growing on minimal and rich media. *J. Bacteriol.* **1999**, *181*, 6425-6440.

88 Richmond, C. S.; Glasner, J. D.; Mau, R.; Jin, H.; Blattner, F. R. Genome-wide expression profiling in *Escherichia coli* K-12. *Nucl. Acid Res.* **1999**, *27*, 3821-3839.

89 Oh, M.-K.; Liao, J. C. In *Book of Abstracts, 219th ACS National Meeting, San Francisco, CA, March 26-30, 2000*, 2000, pp BIOT-161.

90 Yi, J.; Li, K.; Draths, K. M.; Frost, J. W. Modulation of Phosphoenolpyruvate Synthase Expression Increases Shikimate Pathway Product Yields in *E. coli. Biotechnol. Prog.* **2002**, *18*, 1141-1148.

91 (a) Hoischen, C.; Kramer, R. Evidence for an Efflux Carrier System Involved in the Secretion of Glutamate by *Corynebacterium glutamicum. Arch. Microbiol.* **1989**, *151*, 342-347. (b) Kramer, R. Secretion of Amino-Acids by Bacteria - Physiology and Mechanism. *FEMS Microbiol. Rev.* **1994**, *13*, 75-93.

92 (a) Bellmann, A.; Vrljic, M.; Patek, M.; Sahm, H.; Kramer, R.; Eggeling, L. Expression Control and Specificity of the Basic Amino Acid Exporter LysE of *Corynebacterium glutamicum. Microbiology-Sgm* **2001**, *147*, 1765-1774. (b) Vrljic, M.; Sahm, H.; Eggeling, L. A New Type of Transporter with a New Type of Cellular Function: L-Lysine Export from *Corynebacterium glutamicum. Mol. Microbiol.* **1996**, *22*, 815-826. (c) Vrljic, M.; Kronemeyer, W.; Sahm, H.; Eggeling, L. Unbalance of L-Lysine Flux in *Corynebacterium glutamicum* and Its Use for the Isolation of Excretion-Defective Mutants. *J. Bacteriol.* **1995**, *177*, 4021-4027.

93 (a) Kennerknecht, N.; Sahm, H.; Yen, M. R.; Patek, M.; Saier, M. H.; Eggeling, L. Export of L-Isoleucine from *Corynebacterium glutamicum*: A Two Gene-Encoded Member of a New Translocator Family. *J. Bacteriol.* **2002**, *184*, 3947-3956. (b) Ebbighausen, H.; Weil, B.; Kramer, R. Isoleucine Excretion in *Corynebacterium-Glutamicum* - Evidence for a Specific Efflux Carrier System. *Appl. Microbiol. Biotech.* **1989**, *31*, 184-190. (c) Hermann, T.; Kramer, R. Mechanism and Regulation of Isoleucine Excretion in *Corynebacterium glutamicum. Appl. Environ. Microbiol.* **1996**, *62*, 3238-3244. (d) Morbach, S.; Sahm, H.; Eggeling, L. L-Isoleucine Production with *Corynebacterium glutamicum*: Further Flux Increase and Limitation of Export. *Appl. Environ. Microbiol.***1996**, *62*, 4345-4351.

94 (a) Simic, P.; Willuhn, J.; Sahm, H.; Eggeling, L. Identification of Glya (Encoding Serine Hydroxymethyltransferase) and Its Use Together with the Exporter ThrE to Increase L-Threonine Accumulation by *Corynebacterium Glutamicum. Appl. Environ. Microbiol.* **2002**, *68*, 3321-3327. (b) Simic, P.; Sahm, H.; Eggeling, L. L-Threonine Export: Use of Peptides to Identify a New Translocator from *Corynebacterium glutamicum. J. Bacteriol.* **2001**, *183*, 5317-5324. (c) Palmieri, L.; Berns, D.; Kramer, R.; Eikmanns, M. Threonine Diffusion and Threonine Transport in *Corynebacterium glutamicum* and Their Role in Threonine Production. *Archives of Microbiology* **1996**, *165*, 48-54.

95 Dassler, T.; Maier, T.; Winterhalter, C.; Bock, A. Identification of a Major Facilitator Protein from *Escherichia Coli* Involved in Efflux of Metabolites of the Cysteine Pathway. *Mol. Microbiol.* **2000**, *36*, 1101-1112.

96 Zakataeva, N. P.; Aleshin, V. V.; Tokmakova, I. L.; Troshin, P. V.; Livshits, V. A. The Novel Transmembrane *Escherichia coli* Proteins Involved in the Amino Acid Efflux. *FEBS Letters* **1999**, *452*, 228-232.

97 (a) Liu, J. Y.; Miller, P. F.; Gosink, M.; Olson, E. R. The Identification of a New Family of Sugar Efflux Pumps in *Escherichia coli. Mol. Microbiol.* **1999**, *31*, 1845-1851. (b) Carole, S.; Pichoff, S.; Bouche, J. P. *Escherichia coli* Gene YdeA Encodes a Major Facilitator Pump Which Exports L-Arabinose and Isopropyl-Beta-D-Thiogalactopyranoside *J. Bacteriol.* **1999**, *181*, 5123-5125.

## CHAPTER 2

### Availability of phosphoenolpyruvate and biosynthesis of DHS from glucose

### Introduction

3-Dehydroshikimic acid (DHS) is an important intermediate in benzene-free syntheses of a variety of industrial chemicals: commodity chemicals (adipic acid,[1] phenol[2]), pseudocommodity chemicals (catechol,[3] hydroquinone,[4] *p*-hydroxybenzoic acid[5]), fine chemicals (vanillin,[6] indigo[7]), and ultrafine chemicals (gallic acid,[8] pyrogallol[8b]). Strategies developed to improve the synthesis of DHS can thus be applied to the microbial synthesis of a wide range of molecules. In addition, DHS can be used as a potent antioxidant.[9]

The key step for the biosynthesis of DHS from glucose is the first step in the common pathway of aromatic amino acid biosynthesis: condensation of phosphoenolpyruvate (PEP) with D-erythrose 4-phosphate (E4P) catalyzed by 3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate (DAHP) synthase (Figure 14). PEP is drived from the Embden-Meyerhof pathway (glycolysis) and E4P is derived from pentose phosphate pathway. Overexpression of DAHP synthase isozymes insensitive to feedback inhibition formed the initial basis for increasing the carbon flow directed into the shikimate pathway.[10] The in vivo availability of PEP and E4P limit the catalytic activity of overexpressed, feedback-insensitive DAHP synthase in *Escherichia coli* K-12.[11] Access to E4P is increased upon overexpression of *tktA*-encoding transketolase.[12] With improved E4P availability, the availability of PEP becomes a key limiting factor in DHS biosynthesis as discussed in Chapter 1.[13]

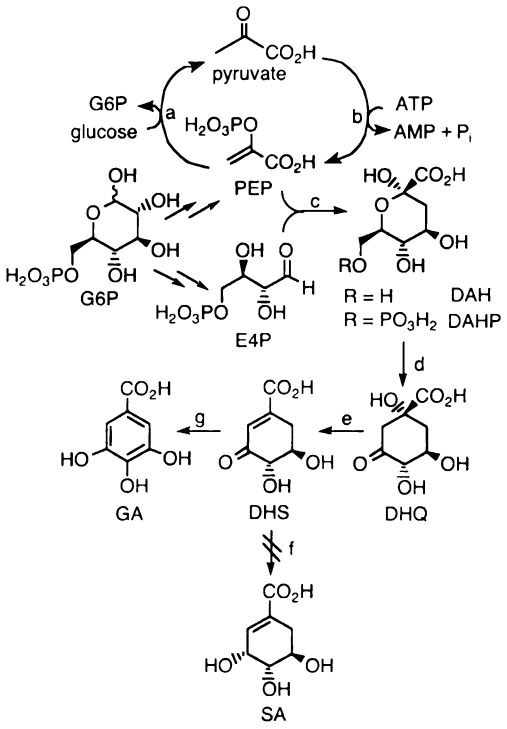



**Figure 14. Synthesis of DHS from D-glucose.** Key: Enzymes (genes): (a) PEP:carbohydrate phosphotransferase system, (b) PEP synthase (*ppsA*), (c) DAHP synthase ($aroF^{FBR}$), (d) DHQ synthase (*aroB*), (e) DHQ dehydratase (*aroD*), (f) shikimate dehydrogenase (*aroE*), (g) uncharacterized. Intermediates (abbreviations); D-glucose 6-phosphate (G6P), phosphoenolpyruvate (PEP), D-erythrose 4-phosphate (E4P), 3-deoxy-D-*arabino*-heptulosonic acid (DAH) 7-phosphate (DAHP), 3-dehydroquinic acid (DHQ), 3-dehydroshikimic acid (DHS), gallic acid (GA), shikimic acid (SA).

PEP is a key intermediate involved in several cellular processes such as phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS), aromatic amino acid biosynthesis, glycolysis, 3-deoxy-D-*manno*-octulosonate biosynthesis and peptidoglycan biosynthesis. When wild-type *E. coli* grows on glucose, the major consumer of PEP is the $PTS^{14}$ system responsible for the simultaneous transport and phosphorylation of glucose (Figure 12, Chapter 1). One molecule of PEP is converted into pyruvic acid for each molecule of glucose transported into the cytoplasm. PTS-generated pyruvic acid is oxidized through the TCA cycle to carbon dioxide and apparently not recycled to PEP under aerobic culture conditions where glucose is the sole

carbon source. As a consequence, stoichiometric analysis (Figure 15A, discussed in detail in the next section) indicated that the maximum theoretical molar yield for DAHP synthesis from glucose without recycling of pyruvic acid to PEP is equal to 0.43 mol/mol.[15] In theory, this 43% maximum yield can be doubled (Figures 15B and 15C, discussed in detail in the next section) if a strain can recycle PTS-generated pyruvic acid back to PEP (the first strategy) or transport glucose without coupling this process to PEP utilization (the second strategy). To evaluate the impact of increased PEP availability on DHS biosynthesis, this study examined these two basic strategies to alleviate PEP limitation in constructs cultured under fed-batch fermentor conditions.

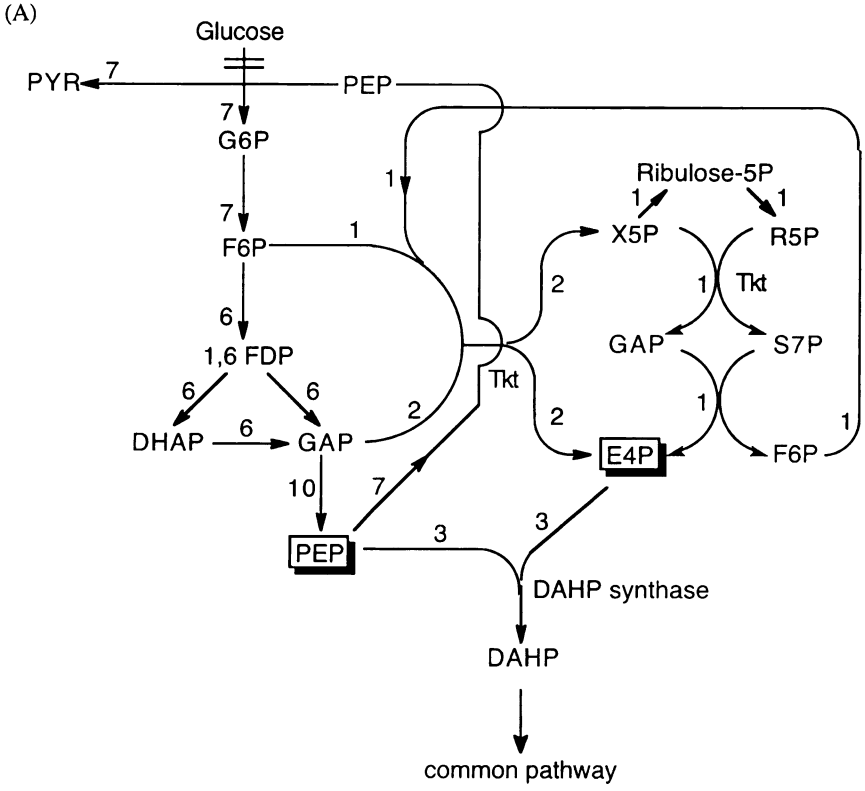
The first strategy employed overexpressed *ppsA*-encoded PEP synthase (Figure 12, Chapter 1) to recycle the pyruvic acid generated by PTS-mediated glucose transport to PEP. PEP synthase converts pyruvic acid to PEP along with conversion of ATP to AMP and one molecule of inorganic phosphate (Figure 12, Chapter 1).[16] In a second strategy, expenditure of PEP during glucose transport was partially or completely avoided by a non-PTS glucose transport mechanism (Figure 12, Chapter 1). Non-PTS glucose transport includes facilitated diffusion mediated by the *Z. mobilis glf*-encoded glucose facilitator[17] and galactose-proton symport mediated by the *E. coli galP*-encoded galactose permease (Figure 12, Chapter 1).[18] Glf is specific for glucose and has an apparent $K_m$ for glucose of approximately 1.1-2.9 mM, consistent with the function of Glf as a low-affinity glucose facilitator.[17] GalP is a member of the major facilitator family (MFS) that can function by solute/$H^+$ symport.[19] Constructs were assembled where PTS-mediated glucose transport was replaced or augmented with heterologous expression of the *Zymomonas mobilis glf*-encoded glucose facilitator protein. Alternatively, PTS-mediated

glucose transport is replaced with the *E. coli galP*-encoded galactose permease. Genomic *E. coli glk*-encoded glucokinase alone or in combination with plasmid-localized *Z. mobilis glk*-encoded glucokinase phosphorylates Glf- or GalP-transported glucose to glucose-6-phosphate using ATP as the phosphor donor. Based on the yield of the same product and byproduct mixture synthesized by *E. coli* constructs cultivated under a uniform set of fermentor-controlled conditions, the impact of strategies employed to alleviate PEP limitation was evaluated.

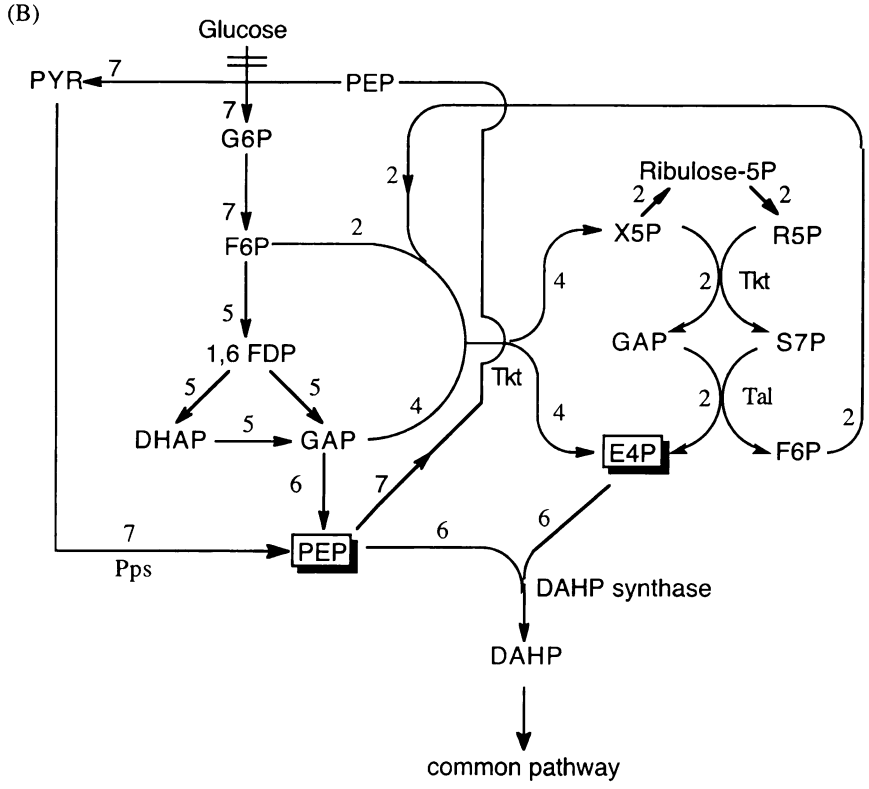
**Theoretic maximum yield calculation**

The theoretic maximum yield is calculated based on the assumption that the branching pathways are blocked and that the carbon flow is directed by the most efficient pathway with minimum loss to carbon dioxide and other metabolites. Under these conditions, the relative flux through each step at the steady state can be calculated by balancing the input (production) and output (consumption) fluxes from each metabolite pool. For instance, the input flux for glyceraldehydes 3-phosphate is 12 mol (6 mol + 6 mol) while the output flux for glyceraldehyde 3-phosphate is also 12 mol (10 mol + 2 mol) in Figure 15A. As shown in Figure 15A, if a strain transports glucose by PTS without overexpression of PEP synthase, the input of 7 mol of glucose can produce 3 mol of DAHP (43% molar yield) and 7 mol of pyruvate, which is further metabolized. To avoid the waste of pyruvate when glucose is used as the carbon source, two approaches are possible: recycling pyruvate to PEP (Figure 15B) or isolation of PTS mutants that use ATP as the phosphate donor (Figure 15C). These approaches all allow 100% theoretical carbon yield, or 86% theoretical molar yield of DAHP (6 mol) from glucose (7 mol).

(A)


Glucose
PYR
7
PEP
7
G6P
7
F6P
6
1,6 FDP
6
6
DHAP
6
GAP
2
10
PEP
1
1
2
Tkt
2
E4P
Ribulose-5P
1
1
X5P
R5P
2
1
Tkt
GAP
S7P
1
2
1
F6P
3
3
DAHP synthase
DAHP
common pathway

(B)


Glucose
PYR
7
PEP
7
G6P
7
F6P
5
1,6 FDP
5
5
DHAP
5
GAP
4
6
7
PEP
7
Pps
2
2
Tkt
4
2
E4P
Ribulose-5P
2
2
X5P
R5P
4
2
Tkt
GAP
S7P
4
2
Tal
2
F6P
6
6
DAHP synthase
DAHP
common pathway

(C)

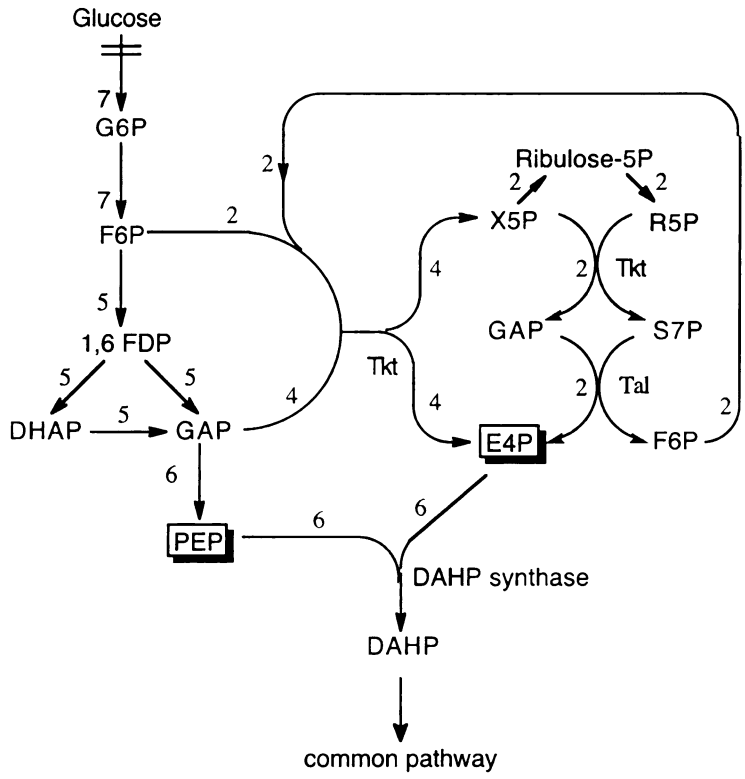



**Figure 15. Theoretical flux distribution for directing carbon into common pathway with glucose as the carbon source.** (A) Without recycling pyruvate to PEP. (B) Recycling pyruvate to PEP using PEP synthase (Pps). (C) Using non-PTS glucose transport. The numbers are the relative fluxes needed to convert 7 mol of glucose into DAHP. Enzymes: Pps, PEP synthase; Tkt, transketolase; Tal, transaldolase. Intermediates (abbreviations): G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; 1,6FDP, 1,6-fructose diphosphate; DHAP, dihydroxyacetone phosphate; GAP, glyceraldehyde 3-phosphate; R5P: ribose 5-phosphate; X5P, xylulose 5- phosphate; S7P, sedoheptulose 7-phosphate; PYR, pyruvate.

The relative flux through each intermediate step is also shown in Figures 15A, 15B and 15C. For instance, 2 (or 4) mol out of 3 (or 6) mol input E4P is coming from the transketolase reaction by coupling Embden-Meyerhof pathway intermediate fructose-6-phosphate and glyceraldehydes 3-phosphate (Figures 15A, 15B and 15C). In reality, this is not likely to happen without overexpression of transketolase.

**Modulation of PEP synthase expression increases shikimate pathway product yields**

**Biocatalyst Design**

*E. coli* KL3/pJY1.216A was constructed to study the impact of phosphoenolpyruvate synthase expression on the synthesis of DHS. *E. coli* KL3[11] carried an *aroE* mutation resulting in the absence of shikimate dehydrogenase (Figure 14). In addition, a second *aroB* locus encoding 3-dehydroquinate synthase was inserted into its *serA* locus encoding 3-phosphoglycerate dehydrogenase. The insertion had two functions: increased 3-dehydroquinate synthase resulting from the two genomic *aroB* loci in *E. coli* KL3 had been previously observed to eliminate formation of DAH as a byproduct when carbon flow directed into the shikimate pathway was increased;[20] disruption of the genomic *serA* locus destroyed the ability of the *E. coli* KL3 to synthesize L-serine and served as a nutritional pressure for maintenance of *serA* localized plasmid in minimal salts medium.

In addition to the *serA* locus, plasmid pJY1.216A (Figure 16C) contained *ppsA*[16] encoding PEP synthase directly behind a tac promoter ($P_{tac}$) and ribosomal binding site. As the $lacI^Q$ gene encoding lac repressor protein was also localized in the plasmid pJY1.216A, the expression of PEP synthase was controlled by the concentration of isopropyl β-D-thioglucopyranoside (IPTG) added to the culture medium. Plasmid pJY1.216A also contained the $aroF^{FBR}$ encoding feedback-insensitive

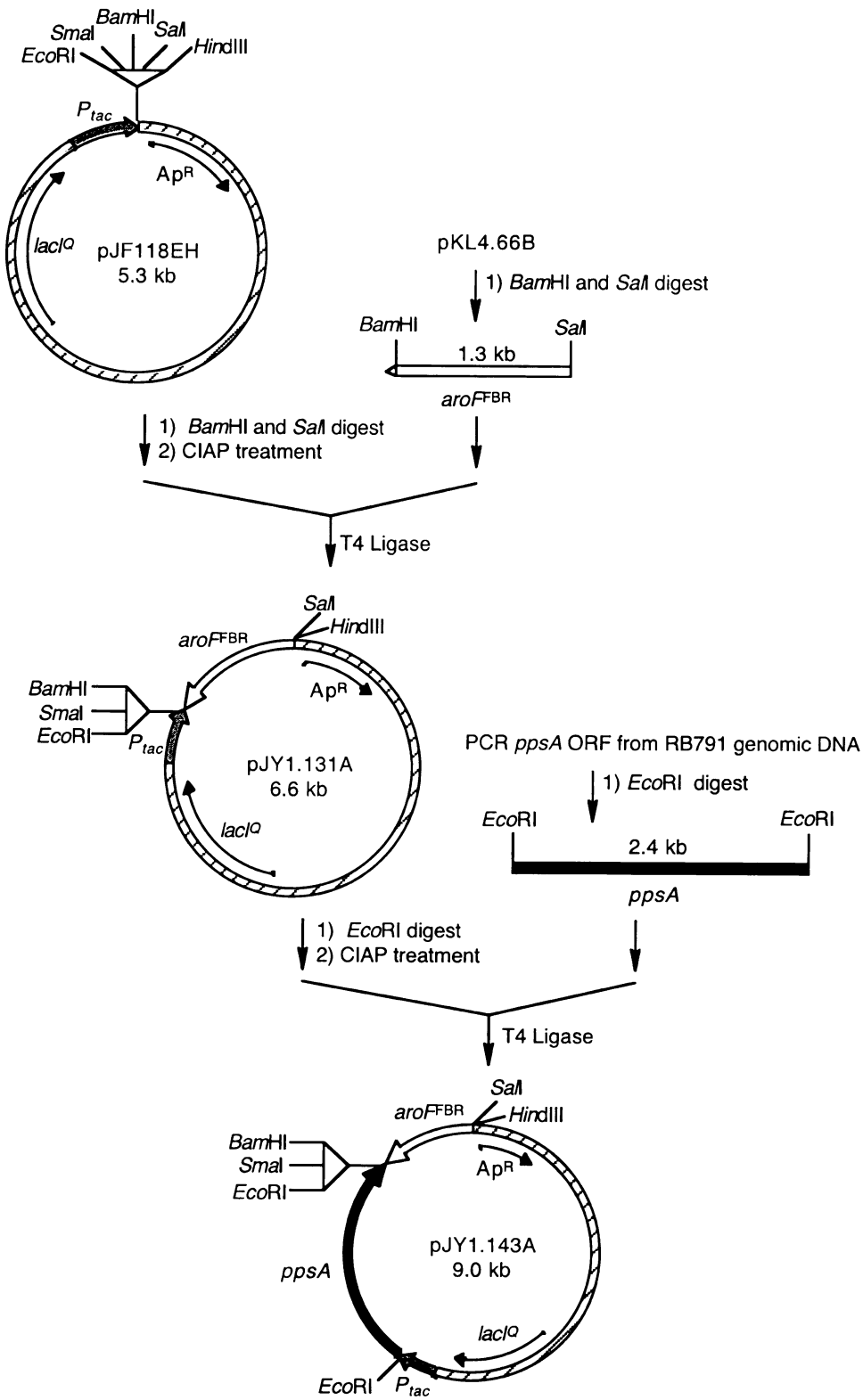



**Figure 16A. Construction of plasmid pJY1.131A and pJY1.143A**

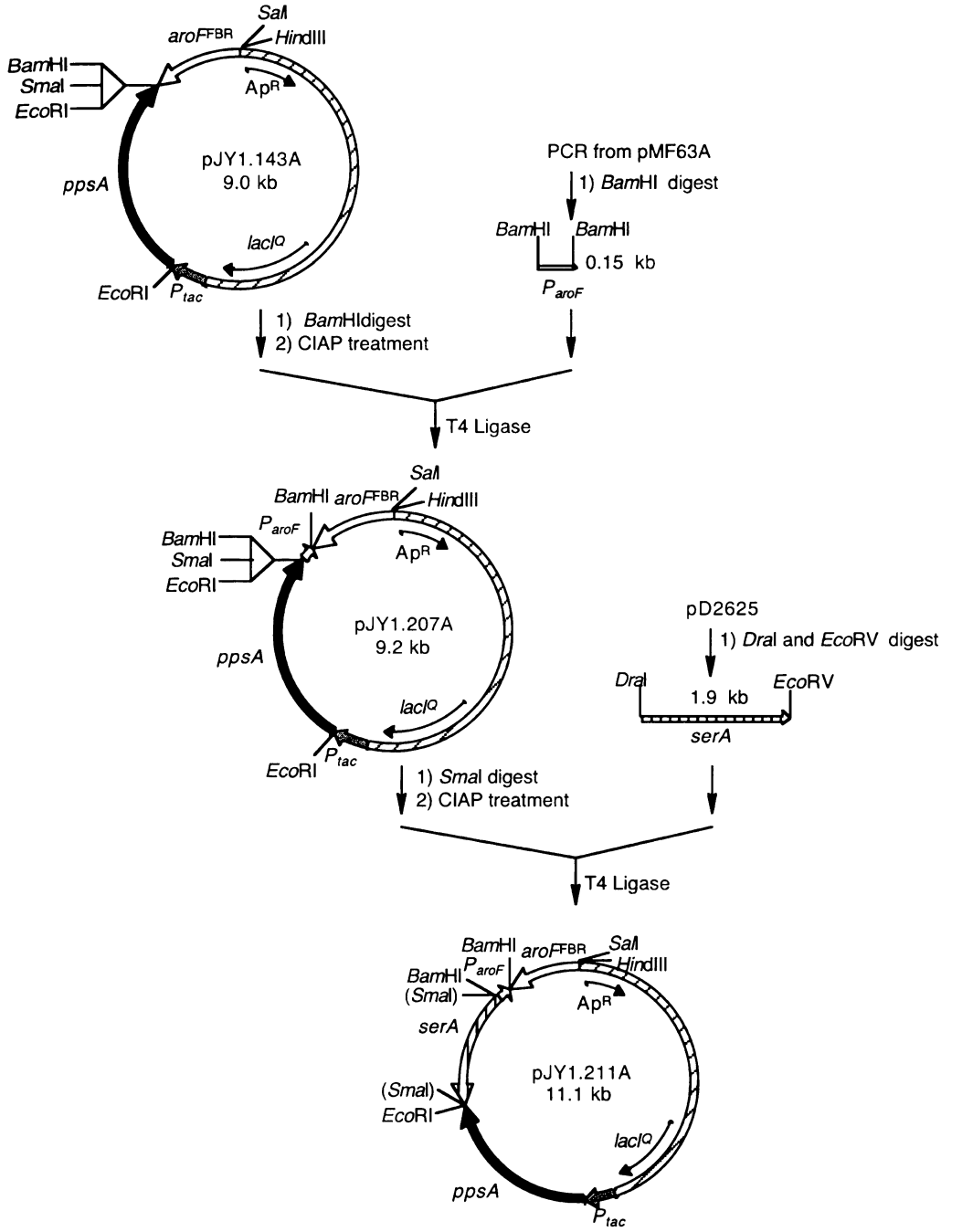



**Figure 16 B. Construction of plasmid pJY1.207A and pJY1.211A**

3-deoxy-D-*arabino*-heptulosonic acid 7-phosphate (DAHP) synthase, an additional copy of promoter region of *aroF* ($P_{aroF}$), and the *tktA* encoding transketolase. In the absence of

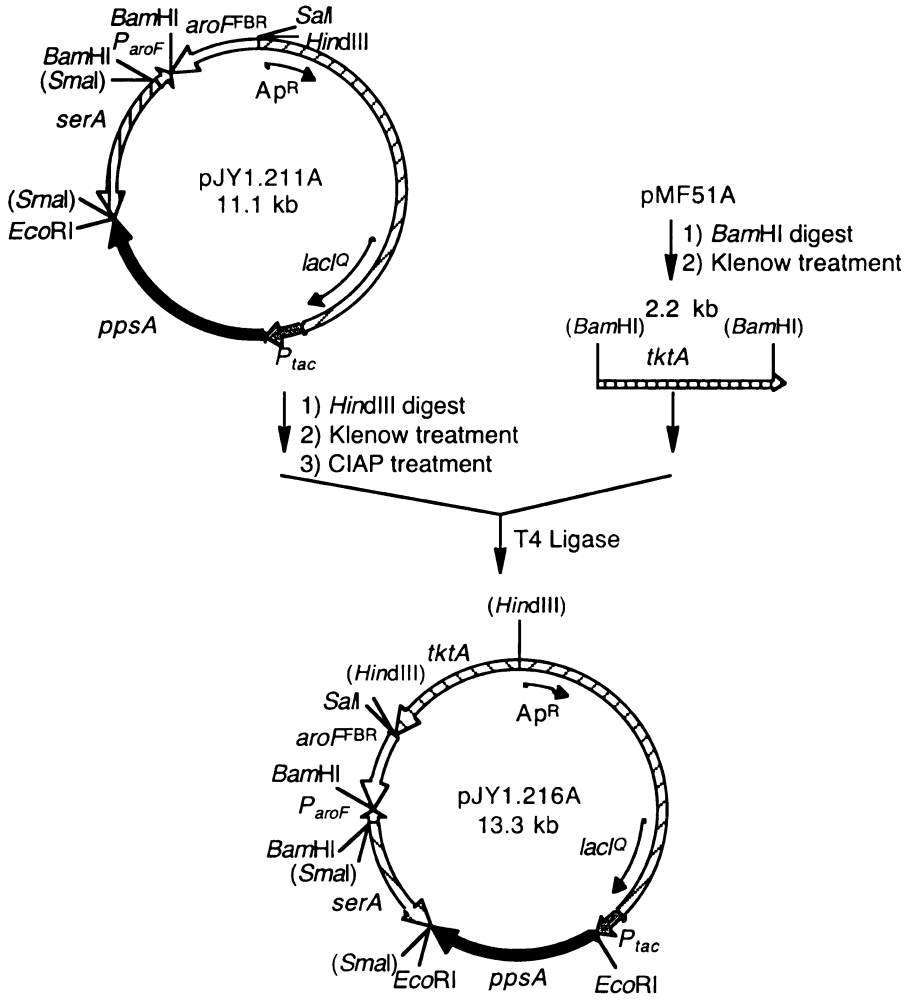



**Figure 16C. Construction of plasmid pJY1.216A**

shikimate dehydrogenase activity, aromatic amino acids and aromatic vitamins had to be added to the culture medium in order for 3-dehydroshikimate-synthesizing constructs to grow. Plasmid-localization of a feedback-insensitive isozyme of DAHP synthase encoded by *aroF*$^{FBR}$ in all 3-dehydroshikimate-synthesizing constructs was thus essential to circumvent feedback inhibition caused by the aromatic amino acid supplements. Plasmid-localized $P_{aroF}$ has three binding sequences for the TyrR repressor protein that represses transcription of *aroF*$^{FBR}$,[21] and its inclusion on a plasmid had been shown to titrate away the cellular supply of TyrR and increased expression of DAHP synthase.[11]

E4P (Figure 14) availability was increased by the expression of plasmid-localized *tktA* encoding transketolase.[12]

**Fermentation conditions**

The impact of expression level of PEP synthase in *E. coli* KL3/pJY1.216A on the yields and concentrations of DHS and shikimate pathway byproducts synthesized under fed-batch fermentor conditions was examined under both glucose-rich conditions and under glucose-limited conditions in a 2.0-L working volume fermentor. A concentration range of 55-170 mM glucose was maintained under glucose-rich conditions while a steady state concentration of approximately 0.2 mM glucose was maintained under glucose-limited conditions. According to the literature, reductions in product yields associated with excessive $CO_2$ generation are minimized under glucose-rich conditions.[22] However, glucose-rich conditions can lead to excessive generation of acetic acid, which is toxic to *E. coli* and many other microbes.[22] Glucose-limited conditions minimize generation of acetic acid but can lead to excessive $CO_2$ generation resulting in lower product yields.[22] Thus both glucose-rich and glucose-limited conditions were employed to find the glucose concentration leading to the synthesis of the highest possible yield of DHS and shikimate pathway byproducts. Temperature was maintained at 36 °C, pH was maintained at 7.0 and dissolved oxygen was maintained at 20% air saturation under both glucose-rich and glucose-limited conditions. All the fermentations were run in duplicate and reported results represent an average of two runs.

**DHS and shikimate pathway byproducts synthesized by *E. coli* KL3/pJY1.216A under glucose-rich conditions**

*E. coli* KL3/pJY1.216A was first tested under glucose-rich conditions. The dissolved oxygen was maintained at 20% air saturation by varying the impeller speed. A glucose concentration of 55-170 mM was maintained by manually adjusting the glucose feeding rate. The fermentation runs stopped at 42 h after inoculation of the culture medium because DHS synthesis did not continue beyond 42 h. To induce the expression of PEP synthase, the indicated quantities of IPTG were added at 18, 24, 30, and 36 h after inoculation of the culture medium. Aliquots of the culture medium at certain time points were removed and concentrations of DHS, 3-dehydroquinic acid, gallic acid, 3-deoxy-D-*arabino*-heptulosonic acid (DAH), and acetic acid were determined by $^1$H NMR using response factors determined with authentic samples of all molecules (Table 1). The specific activities of PEP synthase and DAHP synthase were also measured from aliquots removed from the fermentor at 24, 36, and 42 h after inoculation of the culture medium (Table 2).

At 24 h, PEP synthase specific activities varied from 0.0093 U/mg (Entry 1, Table 2) when no IPTG was added to 0.15 U/mg (Entry 6, Table 2) when 48 mg IPTG was added at 6 h intervals. As expected, PEP synthase activities were higher with addition of more IPTG. For comparison, DAHP synthase specific activities did not vary over more than 2-fold range as a function of IPTG addition, indicating that increased PEP synthase activities did not negatively impact DAHP synthase activities. For all IPTG concentrations examined, both PEP synthase specific activity and DAHP synthase specific activity declined over the course of the fermentation runs after 24 h.

With no IPTG addition and wild-type PEP synthase activity (0.0093 U/mg at 24 h), 52 g/L of DHS was synthesized in 28% (mol/mol) yield from glucose by *E. coli* KL3/pJY1.216A (Entry 1, Table 1) and provided the basis for evaluating the impact on

**Table 1. Concentrations and yields of DHS and byproducts synthesized by *E. coli* KL3/pJY1.216A cultured under glucose-rich conditions.**

| Entry | IPTG addn.[a] | [DHS][b] (g/L) | DHS Yield[c] | [DAH] (g/L) | [DHQ] (g/L) | [GA] (g/L) | Total Yield[d] | Dry cell weight (g/L) | [acetate] (g/L) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 52 | 28% | 3.1 | 7.3 | 7.1 | 37% | 21 | 0.30 |
| 2 | 6.0 | 58 | 31% | 11 | 8.9 | 9.1 | 45% | 21 | 0.30 |
| 3 | 12 | 69 | 35% | 10 | 12 | 13 | 51% | 21 | 0.30 |
| 4 | 18 | 62 | 34% | 11 | 9.8 | 8.0 | 48% | 21 | 0.30 |
| 5 | 24 | 66 | 34% | 9.0 | 9.8 | 10 | 47% | 22 | 0.50 |
| 6 | 48 | 58 | 29% | 12 | 5.6 | 7.0 | 40% | 21 | 0.80 |

[a]Amount (mg) of isopropyl β-D-thioglucopyranoside (IPTG) added at 18, 24, 30, and 36 h. [b] Abbreviations: 3-dehydroshikimic acid (DHS), 3-deoxy-D-*arabino*-heptulosonic acid (DAH), 3-dehydroquinic acid (DHQ), gallic acid (GA). [c] (mol DHS)/(mol glucose consumed). [d](mol DHS + mol DAH + mol DHQ + mol GA)/ (mol glucose consumed).

**Table 2. PEP synthase and DAHP synthase specific activities for *E. coli* KL3/pJY1.216A cultured under glucose-rich conditions.**

| Entry | IPTG addn.[a] | PEP synthase (U/mg)[b] | | | DAHP synthase (U/mg)[c] | | |
|---|---|---|---|---|---|---|---|
| | | 24 h | 36 h | 42 h | 24 h | 36 h | 42 h |
| 1 | 0 | .0093 | .0061 | .0042 | .40 | .32 | .17 |
| 2 | 6.0 | .031 | .018 | .015 | .76 | .27 | .13 |
| 3 | 12 | .069 | .058 | .049 | .53 | .27 | .19 |
| 4 | 18 | .081 | .058 | .046 | .91 | .41 | .28 |
| 5 | 24 | .10 | .069 | .031 | .58 | .44 | .28 |
| 6 | 48 | .15 | .13 | .10 | .87 | .62 | .20 |

[a] Amount of IPTG (mg) added at 18, 24, 30, and 36 h. [b] One unit (U) of PEP (PEP) synthase corresponds to consumption of 1 μmole of pyruvate per min at 30 °C. [c] One unit (U) of DAHP synthase corresponds to formation of 1 μmole of DAHP per min at 37 °C.

the concentration and yield of synthesized DHS attendent with overexpression of PEP synthase. As PEP synthase specific activities increased from 0.0093 U/mg to 0.15 U/mg, higher concentration and yield of DHS was synthesized. Both the concentration and

yield of DHS reached an apparent maximum and then decreased (Entries 1-6, Table 1). With 12 mg of IPTG addition and 7-fold increase in PEP synthase specific activity (0.07 U/mg) at 24 h, 69 g/L of DHS was synthesized in 35% (mol/mol) yield from glucose (Entry 3, Table 1). However, with 48 mg of IPTG addition and 15-fold increase in PEP synthase specific activity (0.15 U/mg) at 24 h, a significantly lower concentration of DHS (58 g/L) was synthesized in reduced yield (29%, mol/mol) from glucose (Entry 6, Table 1) relative to the concentrations and yields of DHS synthesized with 12 mg of IPTG addition (Entry 3, Table 1).

In addition to DHS, shikimate pathway byproducts DAH, 3-dehydroquinic acid, and gallic acid also accumulated in the culture medium (Table 1). During a typical fermentation run of *E. coli* KL3/pJY1.216A (Figure 17), the concentration of DAH reached the maximum at 30 h, the concentration of 3-dehydroquinic acid reached the maximum at 36 h and the concentration of gallic acid reached the maximum at the end of the fermentation run 42 h. The extracellular accumulation of byproducts DAH and 3-dehydroquinic acid (Figure 14) generated from the accumulated DAHP and 3-dehydroquinic acid in vivo indicated that the turnover rate of 3-dehydroquinate synthase and 3-dehydroquinate dehydratase were unable to keep up with the rate of increased carbon flow into the shikimate pathway.[11] Gallic acid (Figure 14) is derived from DHS, although the route for this transformation remains to be discovered. Possibilities include either direct oxidation of DHS or dehydration of this hydroaromatic to protocatechuic acid followed by hydroxylation. As a function of PEP synthase specific activity, the biosynthesized concentrations of DAH, 3-dehydroquinic acid, and gallic acid showed similar trends to that of biosynthesized DHS (Table 1, Figure 17).

The total yield of shikimate pathway products was based on the combined yields of DHS, DAH, 3-dehydroquinic acid, and gallic acid. At the optimum level of PEP synthase expression, 69 g/L DHS, 10 g/L DAH, 12 g/L 3-dehydroquinic acid, and 13 g/L gallic acid were synthesized in 51% total yield (Entry 3, Table 1) by *E. coli* KL3/pJY1.216A under glucose-rich conditions. The maximum concentration of gallic

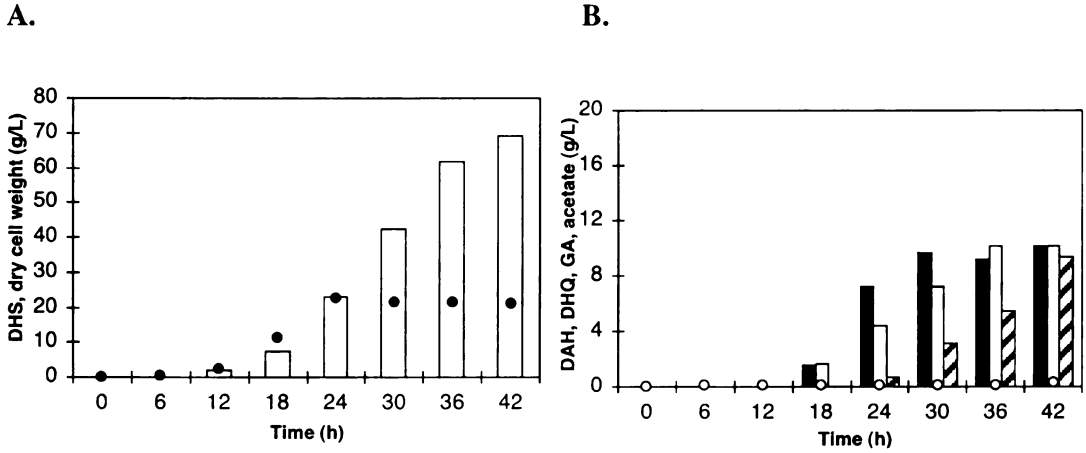



**Figure 17. *E. coli* KL3/pJY1.216A cultured under glucose-rich conditions with 12 mg of IPTG added at 18, 24, 30, and 36 h.** Legend: (A) DHS (open bars), dry cell weight (black circles). (B) 3-deoxy-D-*arabino*-heptulosonic acid (black bars), 3-dehydroquinic acid (open bars), gallic acid (hashed bars), acetic acid (open circles).

acid produced (Entry 3, Table 1) by *E. coli* KL3/pJY1.216A approximates that reported for *E. coli* KL7/pSK6.161 that was specifically designed to synthesize gallic acid with overexpression of both 3-dehydroshikimate dehydratase and protocatechuate hydroxylase.[8]

*E. coli* KL3/pJY1.216A also synthesized relatively low concentrations of acetate (0.3 g/L – 0.8 g/L, Table 1) under glucose-rich conditions. When PEP synthase expression was increased beyond its optimal level for biosynthesis of DHS, acetate accumulation was observed to increase slightly (Entries 5 and 6, Table 1). However, the

increase of acetate concentration did not affect the cell growth as measured by dry cell weight (Entries 1 – 6, Table 1).

The cell mass was approximately 21 g/L at the end of all the fermentation runs (Entries 1 – 6, Table 1). The control profiles of all the fermentation runs were very similar. The same amount of glucose (approximately 1.4 mol) was consumed for all the fermentation runs of *E. coli* KL3/pJY1.216A under glucose-rich conditions, indicating that increasing the expression level of PEP synthase didn't increase the glucose consumption rate for *E. coli* KL3/pJY1.216A.

**DHS and shikimate pathway byproducts synthesized by *E. coli* KL3/pJY1.216A under glucose-limited conditions**

*E. coli* KL3/pJY1.216A was also tested under glucose-limited conditions. Automatic oxygen-sensor-controlled glucose feeding was used to maintain a steady-state low glucose concentration. The dissolved oxygen in the medium was also maintained at 20% air saturation. Under glucose-limited conditions, the fermentation runs stopped at 48 h after inoculation of the culture medium because DHS synthesis did not continue beyond 48 h. The indicated quantities of IPTG were added 18, 24, 30, 36 and 42 h after inoculation of the culture medium. Aliquots of the culture medium at certain time points were removed and the concentrations of DHS and shikimate pathway byproducts were determined by $^{1}$H NMR (Table 3). The specific activities of PEP synthase and DAHP synthase in these aliquots were also measured from the fermentor at 24, 36, and 42 h after inoculation of the culture medium (Table 4).

**Table 3. Concentrations and yields of DHS and byproducts synthesized by *E. coli* KL3/pJY1.216A cultured under glucose-limited conditions.**

| Entry | IPTG addn.[a] | [DHS][b] (g/L) | DHS Yield[c] | [DAH] (g/L) | [DHQ] (g/L) | [GA] (g/L) | Total Yield[d] | Dry cell weight (g/L) | acetate] (g/L) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 36 | 20% | 0 | 4.7 | 4.5 | 25% | 22 | 0.10 |
| 2 | 12 | 52 | 29% | 4.6 | 6.8 | 11 | 41% | 22 | 0.10 |
| 3 | 24 | 48 | 27% | 3.3 | 6.9 | 8.4 | 36% | 22 | 0.10 |

[a] Amount of IPTG (mg) added at 18, 24, 30, and 36 h. [b] Abbreviations: 3-dehydroshikimic acid (DHS), 3-deoxy-D-*arabino*-heptulosonic acid (DAH), 3-dehydroquinic acid (DHQ), gallic acid (GA). [c] (mol DHS)/(mol glucose consumed). [d](mol DHS + mol DAH + mol DHQ + mol GA)/ (mol glucose consumed).

**Table 4. PEP Synthase and DAHP synthase specific activities for *E. coli* KL3/pJY1.216A cultured under glucose-limited conditions.**

| Entry | IPTG addn.[a] | PEP synthase (U/mg)[b] | | | DAHP synthase (U/mg)[c] | | |
|---|---|---|---|---|---|---|---|
| | | 24 h | 36 h | 42 h | 24 h | 36 h | 42 h |
| 1 | 0 | .0050 | .0060 | .0047 | .25 | .16 | .07 |
| 2 | 12 | .067 | .045 | .036 | .65 | .42 | .15 |
| 3 | 24 | .12 | .069 | .12 | .29 | .33 | .13 |

[a] Amount of IPTG (mg) added at 18, 24, 30, and 36 h. [b] One unit (U) of PEP (PEP) synthase corresponds to consumption of 1 μmole of pyruvate per min at 30 °C. [c] One unit (U) of DAHP synthase corresponds to formation of 1 μmole of DAHP per min at 37 °C.

As a function of IPTG addition, the specific activities (Table 4) of both PEP synthase and DAHP synthase measured under glucose-limited conditions were comparable to those measured under glucose-rich conditions (Table 2). However, *E. coli* KL3/pJY1.216A consumed significant less glucose under glucose-limited conditions (approximately 1.2 mol) than under glucose-rich conditions (approximately 1.4 mol). Both the concentrations and the yields of DHS synthesized by *E. coli* KL3/pJY1.216A under glucose-limited conditions were lower than the concentrations and yields synthesized under glucose-rich conditions at comparable levels of PEP synthase *expression*. With no IPTG addition and wild-type PEP synthase activity, *E. coli* KL3/pJY1.216A synthesized 36 g/L of DHS in 20% (mol/mol) yield from glucose (Entry

1, Table 3). DAH was not detected in the medium, although both 3-dehydroquinic acid (4.7 g/L) and gallic acid (4.5 g/L) accumulated, and the total yield was 25%. When PEP synthase expression was increased approximately 10-fold with 12 mg IPTG addition, 52 g/L of DHS was synthesized in 29% yield (mol/mol) from glucose (Entry 2, Table 3). Accumulation of DAH was once again observed and the total yield was 41% (mol/mol). Consistent with the observation under glucose-rich conditions, increasing PEP synthase expression beyond this level (approximately 0.07 U/mg at 24 h) had a negative impact on shikimate pathway product synthesis. When PEP synthase expression was increased approximately 20-fold with 24 mg IPTG addition (Entry 3, Table 3), 48 g/L of DHS was synthesized in 27% (mol/mol) yield from glucose, and the total yield decreased to 36% (mol/mol). As expected, *E. coli* KL3/pJY1.216A synthesized significantly lower acetate concentrations (0.1 g/L, Entries 1 - 3, Table 3) under glucose-limited conditions than the acetate concentrations synthesized under glucose-rich conditions (0.3 - 0.8 g/L, Table 1). The cell mass at the end of the fermentation runs was approximately the same both under glucose-rich conditions and under glucose-limited conditions.

## Altered glucose transport and shikimate pathway product yields in *E. coli*

### Biocatalyst design

The host strain *E. coli* KL3 used for synthesis of DHS was described in the previous section.[11] *E. coli* KL3 utilizes the wild-type phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS) to transport and phosphorylation of glucose. *E. coli*

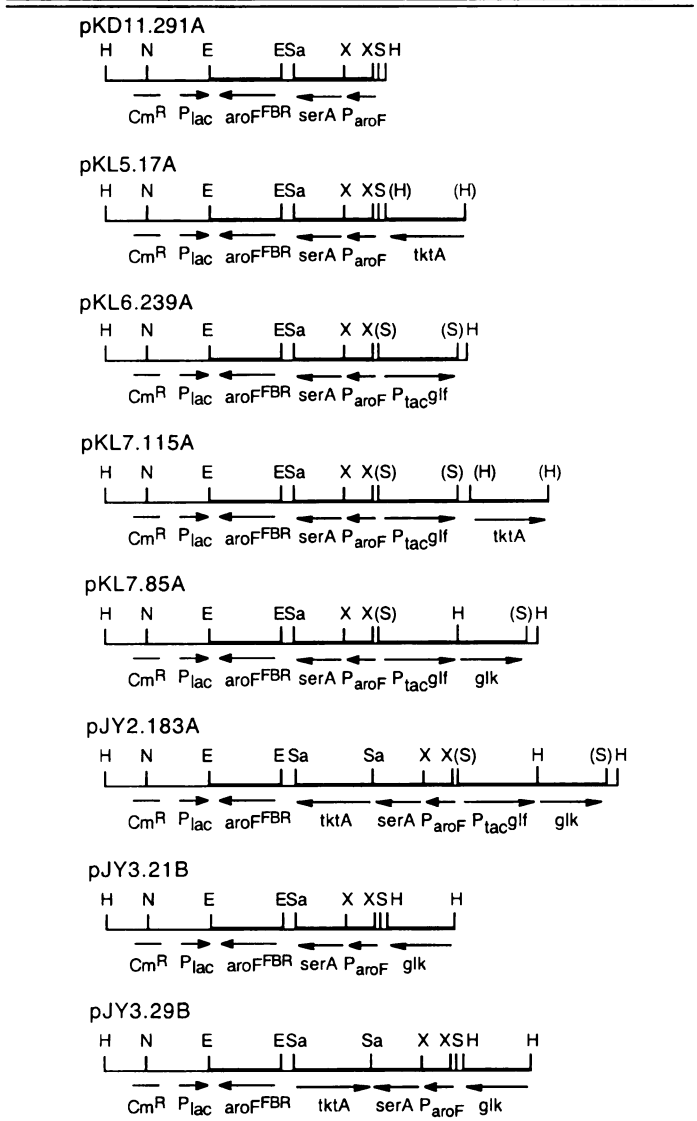



**Figure 18[a]. Restriction enzyme maps of plasmids.**

[a] Restriction enzyme sites are abbreviated as follows: B = *Bam*HI, E = *Eco*RI, H = *Hin*dIII, N = *Nco*I, S = *Sal*I, Sa = *Sac*I, X = *Xba*I. Parentheses indicate that the designed enzyme site has been eliminated.

JY1 (KL3Δ*ptsHIcrr*, devoid of the *ptsH*, *ptsI*, and *crr* genes) was made from *E. coli* KL3 by P1 phage transduction using *E. coli* UE79[23] as a donor strain to inactivate PTS-mediated transport and phosphorylation of glucose. The *ptsH*-encoded protein and *ptsI*-encoded protein are proteins involved in all PTS systems. The *crr*-encoded protein is specific for glucose PTS transport. *E. coli* JY1 was characterized by its growth pattern

on glucose, maltose containing fosfomycin, xylose, and xylose containing cAMP.[24] The isolation of spontaneous mutants that can grow on glucose from *E. coli* strains lacking the *pts* genes has been reported previously.[25] Using a similar approach, cultivation of *E. coli*JY1/pRC55B under fermentor-controlled conditions in M9 glucose minimal medium with aromatic supplementation led to the isolation of a faster-growing variant that was designated *E. coli* JY1.2. Plasmid pRC55B containing *serA* gene enabled the growth of *E. coli* KL3-derived strains in minimal salts medium without serine supplementation. Repetitive subculturing of *E. coli* JY1.2/pRC55B in M9 glucose minimal medium resulted in the isolation of *E. coli* JY1.3. *E. coli* JY1.2 and *E. coli* JY1.3 differed in their growth rates in M9 glucose minimal medium. The inactive PTS in *E. coli* JY1.2 and *E. coli* JY1.3 was confirmed by growth on maltose containing fosfomycin and no growth on mannitol. The GalP dependency for glucose transport in *E. coli* JY1.2 and *E. coli* JY1.3 was confirmed by transduction of a *galP*::Tn*10* mutation to generate strains that were unable to grow on glucose.

The plasmids used to evaluate the impact of different mechanisms for glucose transport on the concentration and yield of synthesized DHS are summarized in Figure 18. Plasmid pKL5.17A contained *tktA*, an additional copy of the $P_{aroF}$ promoter region of *aroF*, and $aroF^{FBR}$. All the other plasmid contained *Z. mobilis glf* encoding the glucose facilitator protein and (or) *Z. mobilis glk* encoding glucokinase in addition to plasmid pKL5.17A. *E. coli* KL3/pKL5.17A transports glucose by wild-type PTS system. *E. coli* JY1.2/pKL5.17A and *E. coli* JY1.3/pKL5.17A transports glucose by galactose permease (GalP) followed by phosphorylation of glucose with native glucokinase. Plasmid

pJY2.183A (Figure 19), which carried both a *Z. mobilis glf* insert and a *Z. mobilis glk*

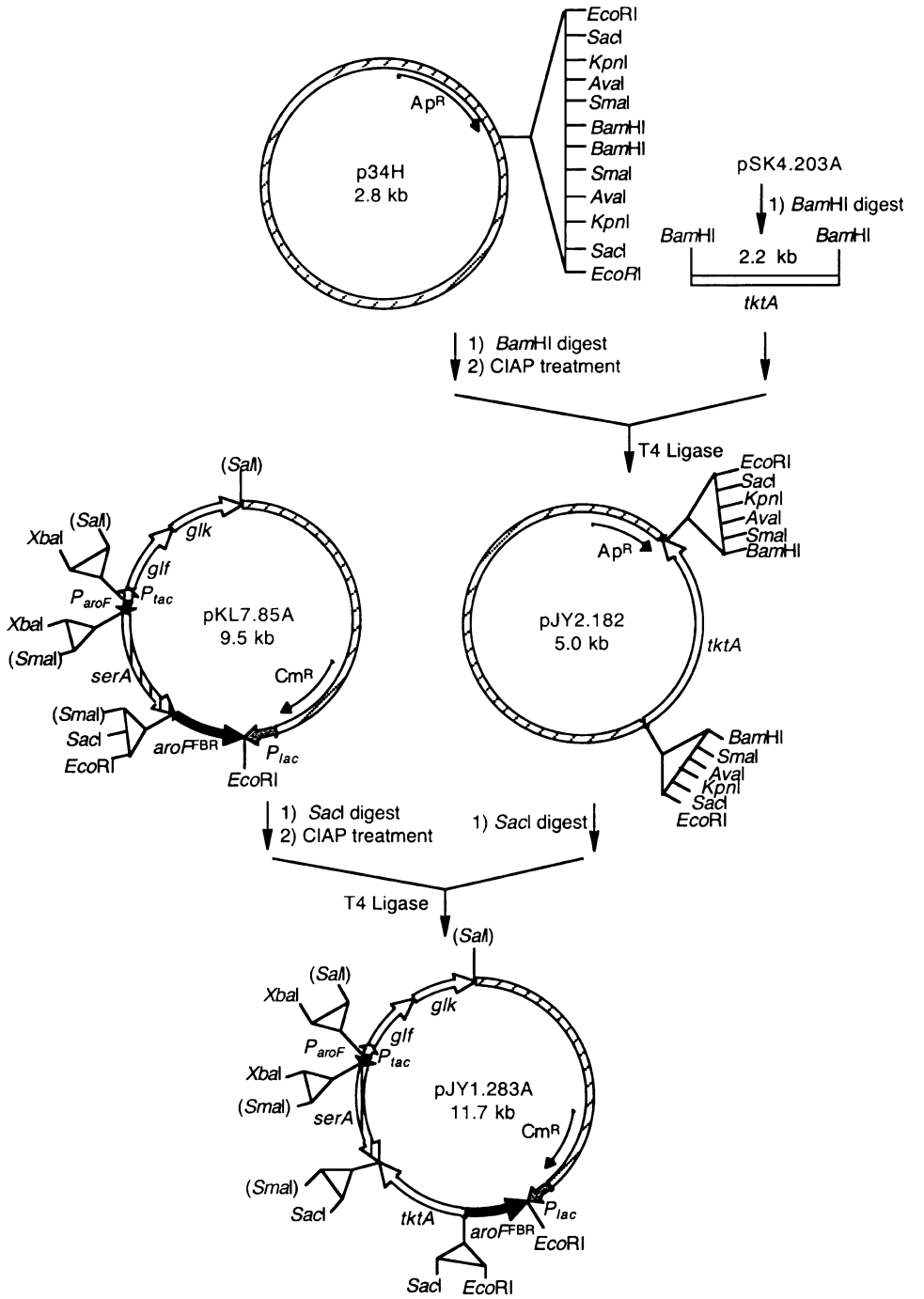



**Figure 19. Construction of plasmid pJY2.182 and pJY2.183A**

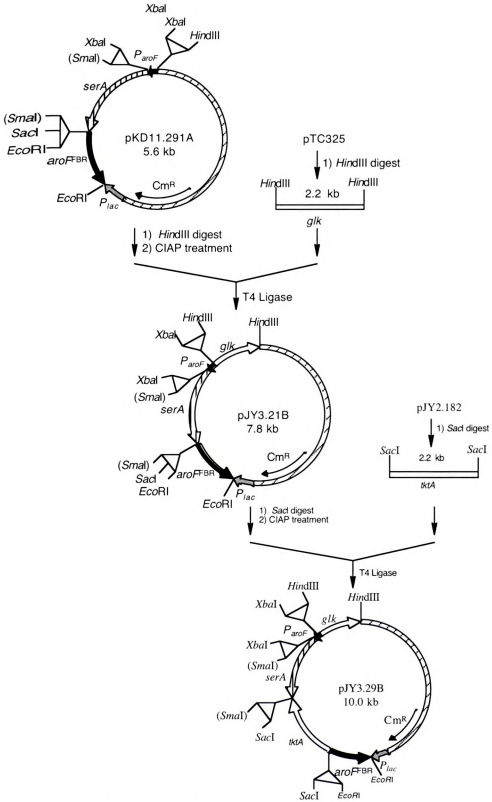



**Figure 20. Construction of plasmid pJY3.21B and pJY3.29B**

insert, was used in *E. coli* JY1/pJY2.183A to transport glucose by Glf followed by phosphorylation of glucose with overexpressed glucokinase (Glk). Plasmid pKL7.115A carried only a *Z. mobilis glf* insert and was used in *E. coli* KL3/pKL7.115A to transport glucose by both PTS and Glf. Plasmid pJY3.29B (Figure 20) carried only a *Z. mobilis glk* insert and was used in *E. coli* JY1.2/pJY3.29B and *E. coli* JY1.3/pJY3.29B to transport glucose by GalP followed by phosphorylation of glucose with overexpressed glucokinase (Glk).

**Concentrations and yields of synthesized DHS**

All constructs were cultured at 1 L scale in a 2.0-L working volume fermentor at pH 7.0 and 36 °C. Dissolved oxygen was maintained at 20% air saturation under both glucose-rich and glucose-limited conditions. Glucose-rich conditions were important for evaluation of both the Glf system and the GalP system since the driving force to transport glucose by facilitated diffusion (Glf) or galactose permease (GalP) included the concentration gradient of glucose between the culture medium and cytoplasm. Aliquots of the culture medium at certain time points were removed and concentrations of DHS, 3-dehydroquinic acid, gallic acid, 3-deoxy-D-*arabino*-heptulosonic acid (DAH), and acetic acid were determined by $^1$H NMR using response factors determined with authentic samples of all molecules (Table 5). The specific activities of DAHP synthase and glucokinase were also measured from aliquots removed from the fermentor at certain time points (Table 6 and Table 7). All the fermentations were run in duplicate and reported results represent an average of two runs.

With native PTS-mediated glucose transport, *E. coli* KL3/pKL5.17A synthesized 49 g/L DHS in 26% yield under both glucose-limited (Entry 1, Table 2) and glucose-rich

conditions (Entry 2, Table 5) and provided the basis for evaluating the impact of different glucose transport mechanisms on concentration and yield of synthesized DHS.

**Table 5. Yields and concentrations of synthesized DHS and shikimate pathway byproducts.**

| Entry | Strain | Host[c]/Plasmid[d] | Time[e] (h) | DHS[f] (g/L) Yield[g] | DAH, DHQ, GA (g/L) | Total Yield[h] | AcOH[f] (g/L) |
|---|---|---|---|---|---|---|---|
| 1[a] | KL3/pKL5.17A | *pts*$^+$ | 48 | (49) 26% | (0.0) (6.0) (4.8) | 33% | 0.1 |
| 2[b] | KL3/pKL5.17A | *pts*$^+$ | 42 | (49) 26% | (0.0) (7.9) (6.1) | 33% | 0.1 |
| 3[a] | KL3/pKL7.115A | *pts*$^+$/*glf* | 48 | (45) 25% | (0.0) (5.1) (6.0) | 32% | 0.0 |
| 4[b] | KL3/pKL7.115A | *pts*$^+$/*glf* | 48 | (63) 27% | (4.1) (7.7) (7.8) | 36% | 0.6 |
| 5[a] | JY1/pJY2.183A | *pts*$^-$/*glf glk* | 48 | (60) 34% | (0.0) (5.5) (8.6) | 41% | 0.2 |
| 6[b] | JY1/pJY2.183A | *pts*$^-$/*glf glk* | 42 | (31) 19% | (0.0) (5.0) (4.4) | 25% | 9.5 |
| 7[b] | JY1.2/pKL5.17A | *pts*$^-$*galP*$^+$ | 96 | (40) 20% | (0.0) (5.4) (5.1) | 25% | 0.0 |
| 8[b] | JY1.2/pJY3.29B | *pts*$^-$*galP*$^+$/*glk* | 84 | (35) 19% | (0.0) (4.2) (5.7) | 24% | 0.8 |
| 9[b] | JY1.3/pKL5.17A | *pts*$^-$*galP*$^+$ | 60 | (60) 36% | (0.0) (5.7) (7.1) | 43% | 0.0 |
| 10[b] | JY1.3/pJY3.29B | *pts*$^-$*galP*$^+$/*glk* | 60 | (54) 29% | (0.0) (5.2) (13) | 38% | 0.0 |
| 11[a] | JY1.3/pKL5.17A | *pts*$^-$*galP*$^+$ | 60 | (37) 21% | (0.0) (4.1) (3.4) | 25% | 0.0 |

[a] glucose-limited conditions. [b] glucose-rich conditions. [c] in addition to *aroE*353 *serA*::*aroB* (see Table 1). [d] in addition to *aroF*$^{FBR}$, $P_{aroF}$, *tktA*, and *serA* plasmid inserts (see Figure 2). [e] fed-batch fermentor culture times. [f] Abbreviations: DHS (DHS), 3-deoxy-D-*arabino*-heptulosonic acid (DAH), 3-dehydroquinic acid (DHQ), gallic acid (GA), acetic acid (AcOH). [g] (mol DHS)/(mol glucose consumed). [h] (mol DHS + mol DAH + mol DHQ + mol GA)/(mol glucose consumed).

Accumulation of DAH was not observed and the total yields of DHS and byproducts 3-dehydroquinic and gallic acids synthesized by *E. coli* KL3/pKL5.17A (33%) were also the same under both glucose-limited (Entry 1, Table 5) and glucose-rich conditions (Entry 2, Table 5). Unlike *E. coli* KL3/pJY1.216A, *E. coli* KL3/pKL5.17A consumed the same amount of glucose (approximately 1.4 mol) under both glucose-rich and glucose-limited conditions.

*E. coli* KL3/pKL7.115A was constructed to transport glucose by both PTS and Glf-mediated facilitated diffusion. Under glucose-rich conditions, *E. coli* KL3/pKL7.115A consumed significantly more glucose (approximately 1.8 mol) relative to *E. coli* KL3/pKL5.17A (approximately 1.4 mol). Under glucose-limited conditions, *E.*

*coli* KL3/pKL7.115A consumed the same amount of glucose (approximately 1.8 mol) as *E. coli* KL3/pKL5.17A, indicating that the outside glucose concentration had a significant impact on the consumption of glucose by Glf-mediated glucose transport. Consistent with more glucose consumption, *E. coli* KL3/pKL7.115A synthesized significantly more DHS (63 g/L) in almost the same yield (27%) from glucose under glucose-rich conditions (Entry 4, Table 5) relative to *E. coli* KL3/pKL5.17A (Entry 2, Table 5). The increase in the total yield of DHS and shikimate pathway byproducts by *E. coli* KL3/pKL7.115A (36%) was primarily due to formation of DAH. By contrast, *E. coli* KL3/pKL7.115A under glucose-limited conditions synthesized the same amount of DHS and shikimate pathway byproducts (Entry 3, Table 5) as *E. coli* KL3/pKL5.17A (Entry 1, Table 5), suggesting that Glf may not transport glucose under glucose-limited conditions when PTS is available.

*E. coli* JY1/pKL7.115A was constructed to transport glucose by only Glf-mediated facilitated diffusion. However, *E. coli* JY1/pKL7.115A grew too slow to be used in the high-density culture fermentations where glucose was the carbon source. In *E. coli* JY1/pKL7.115A, the glucose transported inside the cytoplasm via Glf-mediated facilitated diffusion was phosphorylated by *E. coli* chromosomal *glk*-encoded glucokinase. Heterologous overexpression of *Z. mobilis* glucokinase in addition to *E. coli* native glucokinase led to normal growth rates of *E. coli* JY1/pJY2.183A in minimal salts medium where glucose was the carbon source. *E. coli* JY1/pJY2.183A under glucose-limited conditions synthesized significantly more DHS (60 g/L) in higher yield (34%) from glucose (Entry 5, Table 5) relative to *E. coli* KL3/pKL5.17A (Entry 1, Table 5). However, *E. coli* JY1/pJY2.183A under glucose-rich conditions (Entry 6, Table 5)

led to a very steep decline in the synthesis of DHS. Acetic acid, which increased in concentration throughout the course of the fermentation run of *E. coli* JY1/pJY2.183A, reached a final concentration of 9.5 g/L (Entry 6, Table 5). The accumulation of acetic acid appeared to inhibit the cell growth as indicated by the decreased dry cell weight (15 g/L compared to 22 g/L for all the other fermentation runs) at the end of the fermentation run.

*E. coli* JY1.2 and the more rapidly growing *E. coli* JY1.3 was constructed to transport glucose by only GalP-mediated $H^+$ symport. The glucose transported inside the cytoplasm via GalP-mediated $H^+$ symport was phosphorylated by *E. coli* chromosomal *glk*-encoded glucokinase. *E. coli* JY1.2/pKL5.17A synthesized less DHS (40 g/L) in lower yield (20%) from glucose (Entry 7, Table 5) relative to *E. coli* KL3/pKL5.17A under glucose-rich conditions (Entry 2, Table 5). By contrast, the more rapidly growing *E. coli* JY1.3/pKL5.17A (Entry 9, Table 5) synthesized more DHS (60 g/L) in higher yield (36%) relative to *E. coli* KL3/pKL5.17A under glucose-rich conditions (Entry 2, Table 5). However, *E. coli* JY1.3/pKL5.17A under glucose-limited conditions (Entry 11, Table 5) led to a very steep decline in the synthesis of DHS. Although the *Z. mobilis glk* plasmid insert had a major impact on the synthesis of DHS in *E. coli* JY1/pJY2.183A, a *Z. mobilis glk* plasmid insert in *E. coli* JY1.2/pJY3.29B (Entry 8, Table 5) and *E. coli* JY1.3/pJY3.29B (Entry 10, Table 5), respectively, did not lead to a higher concentration and yield of biosynthesized DHS relative to JY1.2/pKL5.17A (Entry 7, Table 5) and JY1.3/pJY5.17A (Entry 9, Table 5) under glucose-rich conditions.

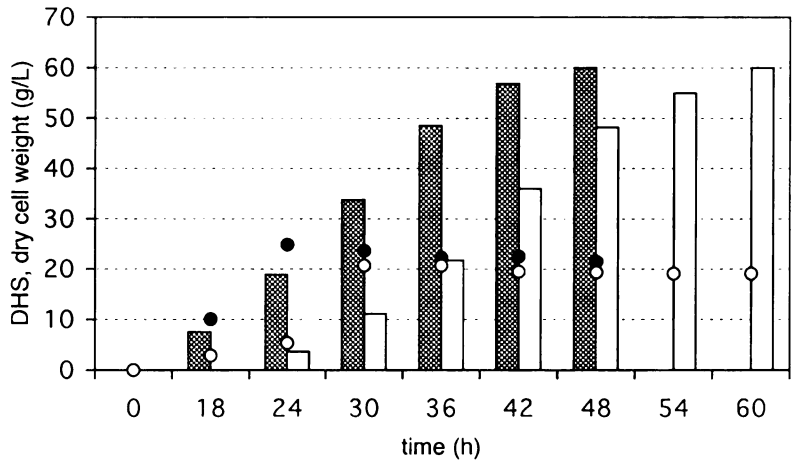



**Figure 21. Growth and synthesis of DHS during cultivation of *E. coli* JY1/pJY2.183A under glucose-limited conditions (gray bars) and *E. coli* JY1.3/pKL5.17A under glucose-rich conditions (open bars).** Dry cell weight formed during growth of *E. coli* JY1/pJY2.183A (black circles) and *E. coli* JY1.3/pKL5.17A (open circles).

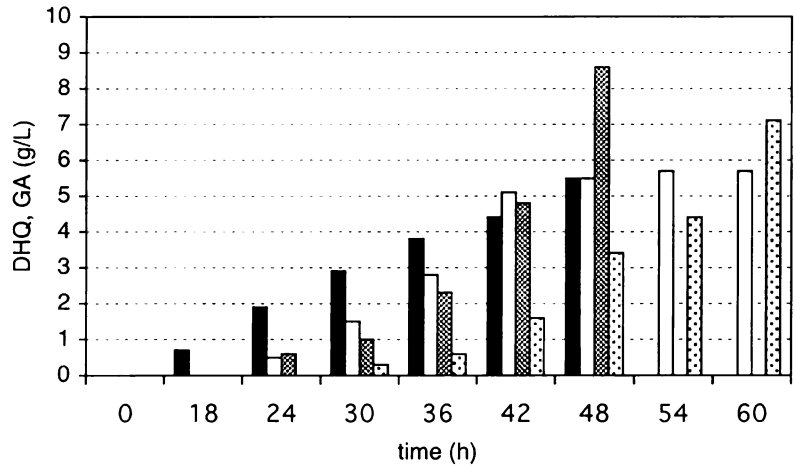



**Figure 22. Synthesis of shikimate pathway byproducts during cultivation of *E. coli* JY1/pJY2.183A under glucose-limited conditions and *E. coli* JY1.3/pKL5.17A under glucose-rich conditions.** 3-Dehydroquinic acid (black bars) and gallic acid (gray bars) synthesized by *E. coli* JY1/pJY2.183A. 3-Dehydroquinic acid (open bars) and gallic acid (stipled bars) synthesized by *E. coli* JY1.3/pKL5.17A.

Heterologous expression of *Z. mobilis glf* and *glk* in *E. coli* JY1/pJY2.183A under glucose-limited conditions and utilizing GalP in *E. coli* JY1.3/pKL5.17A under glucose-rich conditions led to the highest titers and yields of biosynthesized DHS in this study (Figure 21). One major difference between those two fermentation runs was the

cultivation time. *E. coli* JY1/pJY2.183A reached the stationary phase of its growth at approximately 24 h while *E. coli* JY1.3/pKL5.17A reached the stationary phase of its growth at approximately 30 h (Figure 21). *E. coli* JY1/pJY2.183A was cultivated for 48 h until the synthesis of DHS stopped while *E. coli* JY1.3/pKL5.17A was cultivated for 60 h until the synthesis of DHS stopped. The concentrations of both 3-dehydroquinic acid and gallic acid reached the maximum at the end of the fermentation runs of both *E. coli* JY1/pJY2.183A and *E. coli* JY1.3/pKL5.17A (Figure 22).

**DAHP synthase and glucokinase expression**

The specific activity of DAHP synthase varied widely from 0.01 U/mg to 0.98 U/mg as a function of the employed glucose transport system, culture conditions and cultivation times. Under glucose-rich conditions, higher DAHP synthase specific activities (10-26-fold) were observed for PTS-utilizing *E. coli* KL3/pKL5.17A (Entry 2, Table 6) relative to glucose-limited conditions (Entry 1, Table 6). In contrast, under glucose-rich conditions, lower DAHP synthase specific activities were observed for PTS- and Glf- utilizing *E. coli* KL3/pKL7.115A (Entry 3 versus Entry 4, Table 6) and Glf-utilizing *E. coli* JY1/pJY2.183A (Entry 5 versus Entry 6, Table 6) relative to glucose-limited conditions. DAHP synthase specific activities declined after 12 h over the course of the fermentation runs of both PTS-utilizing *E. coli* KL3/pKL5.17A (Entry 2, Table 6) and PTS- and Glf-utilizing *E. coli* KL3/pKL7.115A (Entry 3, Table 6). However, DAHP synthase specific activities were stable for Glf-utilizing *E. coli* JY1/pJY2.183A under glucose-limited conditions (Entry 5, Table 6) and even increasing over the course of the fermentation runs of GalP-utilizing constructs under glucose-rich conditions (Entries 7-9,

Table 6). No correlation was observed between increased synthesized DHS and increased DAHP synthase specific activities.

**Table 6. DAHP synthase specific activities.**

| Entry | Construct | Host[c]/Plasmid[d] | DAHP Synthase Specific Activity (U/mg) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 12 h | 24 h | 36 h | 48 h | 60 h | 72 h |
| 1[a] | KL3/pKL5.17A | *pts*$^+$ | 0.037 | 0.031 | 0.033 | 0.025 | | |
| 2[b] | KL3/pKL5.17A | *pts*$^+$ | 0.98 | 0.63 | 0.30 | 0.23 | | |
| 3[a] | KL3/pKL7.115A | *pts*$^+$/*glf* | 0.52 | 0.53 | 0.26 | 0.11 | | |
| 4[b] | KL3/pKL7.115A | *pts*$^+$/*glf* | 0.12 | 0.64 | 0.15 | 0.01 | | |
| 5[a] | JY1/pJY2.183A | *pts*$^-$/*glf glk* | 0.29 | 0.21 | 0.20 | 0.25 | | |
| 6[b] | JY1/pJY2.183A | *pts*$^-$/*glf glk* | 0.11 | 0.11 | 0.089 | 0.022 | | |
| 7[b] | JY1.2/pKL5.17A | *pts*$^-$*galP*$^+$ | | | | 0.02 | 0.08 | 0.09 |
| 8[b] | JY1.2/pJY3.29B | *pts*$^-$*galP*$^+$/*glk* | | | 0.09 | 0.08 | 0.15 | 0.20 |
| 9[b] | JY1.3/pKL5.17A | *pts*$^-$*galP*$^+$ | | 0.10 | 0.10 | 0.12 | 0.20 | |

[a] glucose-limited conditions. [b] glucose-rich conditions. [c] in addition to *aroE*353 *serA*::*aroB* (see Table 1). [d] in addition to *aroF*$^{FBR}$, $P_{aroF}$, *tktA*, and *serA* plasmid inserts (see Figure 18).

**Table 7. Glucokinase specific activities.**

| Entry | Construct | Host[c]/Plasmid[d] | Glucokinase Specific Activity (U/mg) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 12 h | 24 h | 36 h | 48 h | 60 h | 72 h |
| 5[a] | JY1/pJY2.183A | *pts*$^-$/*glf glk* | 0.30 | 0.30 | 0.30 | 0.30 | | |
| 6[b] | JY1/pJY2.183A | *pts*$^-$/*glf glk* | 1.0 | 1.0 | 1.2 | 1.0 | | |
| 7[b] | JY1.2/pKL5.17A | *pts*$^-$*galP*$^+$ | | | 0.04 | 0.06 | 0.07 | 0.05 |
| 8[b] | JY1.2/pJY3.29B | *pts*$^-$*galP*$^+$/*glk* | | | 0.08 | 0.19 | 0.16 | 0.16 |
| 9[b] | JY1.3/pKL5.17A | *pts*$^-$*galP*$^+$ | | 0.03 | 0.07 | 0.05 | | |
| 10[b] | JY1.3/pJY3.29B | *pts*$^-$*galP*$^+$/*glk* | | 0.11 | 0.15 | 0.10 | | |

[a] glucose-limited conditions. [b] glucose-rich conditions. [c] in addition to *aroE*353 *serA*::*aroB* (see Table 1). [d] in addition to *aroF*$^{FBR}$, $P_{aroF}$, *tktA*, and *serA* plasmid inserts (see Figure 18).

Glucokinase specific activities were stable over the course of all the fermentation runs (Entries 5-10, Table 7). Overexpression of glucokinase in Glf-utilizing *E. coli* JY1/pJY2.183A was a critical factor to enable fast growth on glucose minimal salts medium. The glucokinase specific activity for *E. coli* JY1/pJY2.183A was six-fold higher under glucose-limited conditions (Entry 5, Table 7) and approximately twenty-fold higher under glucose-rich conditions (Entry 6, Table 7) than the average native *E. coli* glucokinase specific activity of 0.05 U/mg (Entry 9, Table 7). For GalP-utilizing

constructs, glucokinase specific activity was observed to be the same for both slow-growing *E. coli* JY1.2/pKL5.17A (Entry 7, Table 7) and faster growing *E. coli* JY1.3/pKL5.17A (Entry 9, Table 7) under glucose-rich conditions, indicating that the reason for the different growth rate between *E. coli* JY1.2 and *E. coli* JY1.3 was probably due to different expression levels of GalP between *E. coli* JY1.2 and *E. coli* JY1.3. The *Z. mobilis glk* insert in *E. coli* JY1.2/pJY3.29B led to a two- to three-fold increase in the specific activity of glucokinase relative to *E. coli* JY1.2/pKL5.17A (Entry 8 versus Entry 7, Table 7). The similar increase in glucokinase specific activity was observed for *E. coli* JY1.3/pJY3.29B relative to *E. coli* JY1.3/pKL5.17A (Entry 10 versus Entry 9, Table 7). However, increased glucokinase activity had little impact on the biosynthesized DHS for GalP-utilizing constructs (Entry 7-10, Table 5).

**Discussion**

Metabolic engineering to design and construct microorganisms suitable for the production of shikimate pathway products requires control of a complicated metabolic network. 3-Deoxy-D-*arabino*-heptulosonic acid 7-phosphate (DAHP) synthase, the first enzyme in the shikimate pathway, is always the key target for manipulation. Overexpression of DAHP synthase isozymes insensitive to feedback inhibition formed the initial basis for increasing the carbon flow directed into the shikimate pathway and increasing the yield of natural products synthesized from glucose via this pathway. With overexpression of a feedback insensitive isozyme of DAHP synthase,[10] Frost and coworkers suggested that the availability of E4P was a key factor limiting DAHP synthase activity.[11, 12] Overexpression of transketolase under shake flask[11] and fermentor-

controlled[12] culture conditions led to significant increase in the concentrations and yields of shikimate pathway products. With the increased availability of E4P, the availability of DAHP synthase's other substrate, PEP, remains as a target for manipulation.

As discussed in the introduction chapter, previous examination of altered glucose transport and overexpressed PEP synthase under shake flask conditions have many disadvantages. Fermentor-controlled cultivation was employed throughout this examination of microbial synthesis of DHS from glucose. An accurate evaluation of yields was thus possible when microbial growth and synthesis of DHS and shikimate pathway byproducts occurred during a single fermentation run.

Glucose-limited or glucose-rich concentrations were maintained in the fermentation. The glucose-rich conditions were previously avoided because of reports of acetate accumulation and low-yield conversion.[22] However, Frost and coworkers previously investigated the production of shikimic acid under both glucose-rich conditions and glucose-limited conditions and observed that concentrations of biosynthesized shikimic acid were significantly higher under glucose-rich conditions.[26] Thus the glucose-rich conditions were employed in this study and were the major difference between this study and earlier study of the impact of overexpression of transketolase and DAHP synthase on the biosynthesis of DHS.[11]

**Increased flux to shikimate pathway product by PEP synthase overexpression**

Overexpression of PEP synthase in *E. coli* KL3/pJY1.216A led to substantial increases in the total yields of DHS and other shikimate pathway byproducts (Table 1 and Table 3) under both glucose-limited and glucose-rich conditions. The increase in yield of

DHS and shikimate pathway byproducts (Table 1) synthesized by *E. coli* KL3/pJY1.216A overexpressing PEP synthase under glucose-rich conditions was particularly significant. As PEP synthase specific activity was almost the same under both glucose-rich and glucose-limited conditions with IPTG addition (Table 2 and Table 4), the yield difference between glucose-rich and glucose-limited conditions with PEP synthase overexpression may due to the different pyruvic acid availability for PEP synthase under different conditions. Under glucose-rich conditions, *E. coli* KL3/pJY1.216A was observed to consume glucose more rapidly (1.4 mol in 42 h) relative to glucose-limited conditions (1.2 mol in 48 h). Due to PTS-modulated glucose transport in *E. coli* KL3/pJY1.216A, the increased glycolysis flux measured as the glucose consumption rate over the course of the fermentation runs under glucose-rich conditions should lead to increased in vivo concentrations of pyruvic acid under glucose-rich conditions. At 51% (mol/mol), the highest total yield (Entry 3, Table 1) of DHS, DAH, 3-dehydroquinic acid, and gallic acid synthesized by *E. coli* KL3/pJY1.216A was well above the maximum theoretical yield of 43% (mol/mol) expected in the absence of recycling of pyruvate to PEP (Figure 15A). The 51% yield, although still well below the maximum theoretical yield of 86% (mol/mol) expected in the presence of recycling of all the PEP synthase-generated pyruvate back to PEP, indicated that the limits imposed on synthesis of DHS by PTS-mediated glucose transport have been successfully circumvented. In the theoretical calculations, neither 86% nor 43% yield takes into consideration the amount of glucose that must be converted into cell biomass or converted into energy to assemble cell mass.

An optimal level of PEP synthase expression (approximately 0.07 U/mg at 24 h) was observed to result in the highest titer and yield of DHS and shikimate pathway byproducts synthesized from glucose by *E. coli* KL3/pJY1.216A under both glucose-rich and glucose-limited conditions. Exceeding this optimal level of PEP synthase expression led to decreases in both the yield and concentration of DHS and shikimate pathway byproducts synthesized from glucose (Table 1 and Table 3). With high PEP synthase expression, it is very likely that the cell's ability to generate energy for cellular processes becomes limited by reduced pyruvic acid availability when the glucose consumption stayed the same for all the fermentation runs under glucose-rich conditions or under glucose-limited conditions. However, the 0.15 U/mg of the highest PEP synthase specific activity observed in this study was much lower than the previous report[27] of 0.29 U/mg of PEP synthase specific activity that only stimulated oxygen consumption and 0.60 U/mg of PEP synthase specific activity that stimulated both oxygen and glucose consumption by approximately 2-fold.

An increase in acetate accumulation (Table 1 versus Table 3) was observed under glucose-rich conditions for *E. coli* KL3/pJY1.216A. However, these concentrations of acetate did not negatively impact microbial growth based on the dry cell weight of *E. coli* KL3/pJY1.216A (Table 1 versus Table 3). Aerobic acetate production is a manifestation of the imbalance between glucose uptake and the demands for both biosynthesis and energy production.[28] The most common arguments for the cause of such imbalance are that the glucose uptake rate is improperly controlled and that the activity of the tricarboxylic acid (TCA) cycle is limiting.[29] So the absence of pyruvate and the modest concentrations of acetate generated by *E. coli* KL3/pJY1.216A can likely be explained by

the properly controlled glucose uptake rate and the activated tricarboxylic acid cycle throughout the fermentation.

**Increased flux to shikimate pathway product by altered glucose transport**

All of the previous examinations of different transport systems have been reviewed in the introduction chapter and there was no direct comparison between GalP-mediated glucose transport and Glf-mediated glucose transport under a common set of culture conditions. In this study, the uniform set of fermentor-controlled as opposed to shake-flask culture conditions were used to maintain glucose, dissolved oxygen concentrations, and pH. Altered glucose transport led to significant increases in the total yields of DHS and shikimate pathway byproducts (Entries 4 and 5, Table 8) relative to the 33% total yield of DHS, 3-dehydroquinic, and gallic acids synthesized by *E. coli* KL3/pKL5.17A utilizing PTS (Entry 1, Table 8). Replacement of the native PTS with Glf-mediated transport and Glk-catalyzed phosphorylation of glucose in *E. coli* JY1/pJY2.183A (Entry 4, Table 8) led to a 41% total yield of DHS, 3-dehydroquinic, and gallic acids under glucose-limited conditions. Replacement of the native PTS with GalP-mediated glucose transport and Glk-catalyzed phosphorylation of glucose in *E. coli* JY1.3/pKL5.17A (Entry 5, Table 8) led to a 43% total yield of DHS, 3-dehydroquinic, and gallic acids under glucose-rich conditions. The results demonstrated that in vivo PEP availability for biosynthesis of shikimate pathway products increased when the PTS was replaced with systems that do not expend PEP during the glucose transport process.

**Table 8. Comparison of the yields and concentrations of synthesized shikimate pathway products and byproducts as a function of the strain and strategy employed to increase PEP availability.**

| Entry | Strain | Glucose transport[c] | Pyruvate recycling[d] | product/byproduct (g/L)[e] | | | | | | Total Yield[g] |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | DAH[f] | DHQ[f] | GA[f] | DHS[f] | QA[f] | SA[f] | |
| 1[b] | KL3/pKL5.17A | PTS | – | 0 | 7.9 | 6.1 | **49** | 0 | 0 | 33 |
| 2[b] | KL3/pJY1.216A | PTS | + | 10 | 12 | 13 | **69** | 0 | 0 | 51 |
| 3[b] | KL3/pKL7.115A | PTS/Glf | – | 4.1 | 7.7 | 7.8 | **63** | 0 | 0 | 36 |
| 4[a] | JY1/pJY2.183A | Glf | – | 0 | 5.5 | 8.6 | **60** | 0 | 0 | 41 |
| 5[b] | JY1.3/pKL5.17A | GalP | – | 0 | 5.7 | 7.1 | **60** | 0 | 0 | 43 |
| 6[b,h] | SP1.1/pKD12.138A | PTS | – | 0 | 0 | 0 | 11 | 4 | **52** | 24 |
| 7[b,h] | SP1.1/pKD15.071B | PTS | + | 0 | 0 | 0 | 16 | 4 | **66** | 29 |
| 8[b,h] | SP1.1/pSC5.112B | PTS/Glf | – | 0 | 0 | 0 | 19 | 6 | **70** | 32 |
| 9[b,h] | SP1.1*pts*/pSC6.090B | Glf | – | 0 | 0 | 0 | 15 | 5 | **71** | 34 |

[a] glucose-rich conditions. [b] glucose-limited conditions. [c] PEP:carbohydrate phosphotransferase (PTS), augmentation of PTS with facilitated diffusion (PTS/Glf), facilitated diffusion (Glf), galactose permease (GalP). [d] native expression (–), amplified expression (+) of phosphoenolpruvate synthase. [e] product targeted for synthesis is in boldface. [f] Abbreviations: 3-deoxy-D-*arabino*-heptulosonic acid (DAH), 3-dehydroquinic acid (DHQ), gallic acid (GA), DHS (DHS), quinic acid (QA), shikimic acid (SA). [g] (mol DAH + mol DHQ + mol GA + mol DHS + mol QA + mol SA)/(mol glucose consumed); Unit: percent. [h] See Ref. 36.

*E. coli* JY1.3/pKL5.17A utilizing GalP (Entry 5, Table 8) without overexpression of *Z. mobilis glk*-encoded glucokinase synthesized approximately the same concentrations of DHS and byproducts 3-dehydroquinic and gallic acids in the same yields as those achieved with *E. coli* JY1/pJY2.183A (Entry 4, Table 8) utilizing Glf with overexpression of *Z. mobilis glk*-encoded glucokinase. However, the slow rates for growth and synthesis of DHS (Figure 21) and shikimate pathway byproducts (Figure 22) is a major disadvantage for GalP-utilizing *E. coli* JY1.3/pKL5.17A relative to Glf-utilizing *E. coli* JY1/pJY2.183A. In addition, the basis for the improvement in growth rate and yields of shikimate pathway products for *E. coli* JY1.3 relative to *E. coli* JY1.2 is unknown and complicates formulation of future metabolic engineering strategies to further improve growth rates and shikimate pathway product yields by GalP system. A

disadvantage of Glf-utilizing *E. coli* JY1/pJY2.183A is that only glucose-limited conditions can be used for the production of DHS. Under glucose-rich conditions, the acetate produced by *E. coli* JY1/pJY2.183A (Entry 6, Table 5) significantly impaired cell growth and metabolic activities and suggested an imbalance between glucose uptake rate and TCA cycle activities.

**Metabolic consequences of PTS inactivation**

The presence of a rapidly metabolizable carbon source such as glucose in the growth medium of a microorganism can inhibit the synthesis of enzymes involved in the metabolism of other carbon-containing compounds. This phenomenon is called catabolic repression. PTS not only functions in glucose transport but also plays an important and complex regulatory role in catabolic repression.[24a, 30] It is expected that a strain without PTS should have impaired catabolic repression and may consume sugar mixtures faster than a wild-type strain. For example, one PTS mutant can assimilate glucose and arabinose simultaneously.[31] Therefore, it would be interesting to determine the growth characteristics of non-PTS strains *E. coli* JY1 and *E. coli* JY1.3 on different carbon sources or sugar mixtures.

**Comparison of altered glucose transport and PEP synthase overexpression**

The strategy of altering glucose transport to increase PEP availability focuses on the competition for PEP molecule by different cellular processes, the same principle used by mutational inactivation of PEP carboxylase[32] and pyruvate kinase.[33] Alternatively, overexpression of PEP synthase recycles PTS-generated and pyruvate kinase-generated

pyruvate back to PEP.[34] The strategy of PEP synthase overexpression has the advantage of recycling not only the PTS generated pyruvate back to PEP but also the pyruvate kinase-generated pyruvate back to PEP. The supply of PEP thus depends on the competition for pyruvate by PEP synthase and the other cellular processes that consume pyruvate. *E. coli* KL3/pJY1.216A, when expressing the optimal specific activity of PEP synthase,[34] synthesized 69 g/L of DHS in a combined yield with DAH, 3-dehydroquinic, and gallic acids of 51% (Entry 2, Table 8) under glucose-rich conditions. This yield is significantly more than the yields of DHS and shikimate pathway byproducts realized with alteration of glucose transport (Entries 3, 4, and 5, Table 8) and may imply that overexpression of PEP synthase is a better strategy for increasing PEP availability. However, the enzyme activities of PEP synthase were optimized in this study while the activities of enzymes involved in altered glucose transport were not optimized. In addition, a number of other factors need to be considered. The acetic acid accumulation with cultivation of Glf-utilizing *E. coli* JY1/pJY2.183A (Entry 6, Table 6) under glucose-rich conditions prevented use of the glucose-rich conditions observed (Entry 2, Table 8) to lead to the highest yields of DHS and shikimate pathway byproducts when PEP synthase expression was amplified. The underlying reason for the faster growth rate and the higher combined yield of DHS, 3-dehydroquinic, and gallic acids synthesized by constructs based on GalP-utilizing *E. coli* JY1.3 relative to *E. coli* JY1.2 (Table 5) remains to be elaborated.

**Comparison of shikimic acid biosynthesis and DHS biosynthesis**

Overexpression of PEP synthase and Glf-mediated facilitated diffusion were also examined in *E. coli* for synthesis of shikimic acid (Entries 6-9, Table 8).[35] GalP-mediated transport was not examined for shikimic acid synthesis because of the uncertain genotype of cells relying on GalP for glucose transport. All shikimate-synthesizing constructs (Entries 6-9, Table 8) contained plasmid-localized $aroF^{FBR}$ encoding feedback-insensitive DAHP synthase, *tktA* encoding transketolase and *aroE* encoding shikimate dehydrogenase. *E. coli* SP1.1/pSC5.112B transported glucose using both Glf and PTS (Entry 8, Table 8). *E. coli* SP1.1*pts*/pSC6.090B transported glucose using only Glf (Entry 9, Table 8). Recycling of PTS-generated pyruvic acid to PEP with overexpression of PEP synthase was tested in *E. coli* SP1.1/pKD15.071B while the plasmid-localized *ppsA* gene was under its native promoter (Entry 7, Table 8). Shikimic acid, DHS and quinic acid were accumulated in the culture medium and the total yield of hydroaromatic products was based on the combined yields of shikimic acid, DHS, and quinic acid (Entries 6-9, Table 8).

Contrary to synthesis of DHS (Entries 1-5, Table 8), replacing (Entry 9, Table 8) or augmenting PTS with Glf (Entry 8, Table 8) were led to higher yields relative to overexpression of PEP synthase (Entry 7, Table 8) during synthesis of shikimic acid. However, this may not be an appropriate comparison because the expression level of PEP synthase was not optimized during the shikimate study. Synthesized yields (29-34%) of shikimic acid and pathway byproducts (Table 8) were consistently lower than the synthesized yields (36-51%) of DHS and pathway byproducts (Table 8) no matter what strategies were employed to increase PEP availability. The variability of the host strain

for synthesis of DHS (*E. coli* KL3[34]) and synthesis of shikimic acid (*E. coli* SP1.1[35]) may contribute to these yield differences. The consumption of NADPH by shikimate dehydrogenase (AroE) may also contribute to the yield differences due to the redox balances inside the cells. Feedback inhibition of shikimate dehydrogenase by shikimic acid may also be limiting yields for synthesis of this hydroaromatic.

**Conclusion**

Overexpression of PEP synthase or alteration of glucose transport by Glf-mediated facilitated diffusion or the GalP galactose permease in *E. coli* constructs led to increased yields of DHS and the combined yields of DHS, 3-dehydroquinic, and gallic acids relative to the native PTS without PEP synthase overexpression. Both strategies were compared under a common set of fermentor-controlled conditions. Overexpression of PEP synthase is currently leading to the synthesis of higher yields of DHS and higher combined yields of DHS and shikimate pathway byproducts relative to alteration of glucose transport. Beyond the optimizations remaining to be made in Glf-utilizing and GalP-utilizing *E. coli* strains for DHS biosynthesis, the general applicability of strategies for increasing PEP availability remained to be tested. The consequences of a given strategy for increasing PEP availability differed as a function of whether DHS or shikimic acid is targeted for synthesis. The respective contribution to this variability arising from the progenitor host *E. coli* strain was investigated in the next section.

## Comparison of KL3*aroKaroL* and SP1.1 for the production of shikimic acid

**Biocatalyst design**

Both shikimate kinase isozymes (Figure 1) encoded by *aroK* and *aroL* were inactivated by successive P1 phage-mediated transductions of *aroL478*::Tn*10* and *aroK*::$Cm^R$ into the appropriate *E. coli* host (KL3 and RB791) to generate *E. coli* constructs (KL3*aroKaroL* and SP1.1).[35] The *aroE* gene localized on plasmid complements genomic mutation of *aroE* in KL3. While ensuring that carbon flow directed into the common pathway did not proceed beyond synthesis of shikimic acid, inactivation of the shikimate kinase also precluded de novo biosynthesis of aromatic amino acids and aromatic vitamins. Growth of all constructs therefore required supplememntation with L-phenylalanine, L-tyrosine, L-trptophan, *p*-hydroxybenzoic acid, *p*-aminobenzoic acid, and 2,3-dihydroxybenzoic acid. These supplements could potentially create a problem in that feedback inhibition of DAHP synthase by aromatic amino acids play a prominent role in controlling carbon flow directed into the common pathway. Accordingly, an isozyme of DAHP synthase encoded by plasmid-localized $aroF^{FBR}$ that was insensitive to feedback inhibition by aromatic amino acids was carried as a plasmid-localized insert. An additional copy of the *aroE* encoding shikimate dehydrogenase, *ppsA* encodign PEP synthase, and *tktA* encoding transketolase, were also localized in the plasmid pKD15.071B[35] to enhance the carbon flow into the shikimate pathway.

**Shikimic acid titers and yields as a function of the progenitor host**

*E. coli* SP1.1/pKD15.071B and KL3*aroKaroL*/pKD15.071B were initially cultured at 1 L scale in a 2.0-L working volume fermentor for 60 h at pH 7.0 and 33 °C. Glucose-rich conditions were employed where the concentration of glucose in the

medium was maintained in the range of 55-170 mM throughout the run. These culture conditions have been identified to suppress equilibration of shikimic acid and quinic acid.[36] Formation of quinic acid results from the reduction of 3-dehydroquinic acid catalyzed by *aroE* encoding shikimate dehydrogenase. The formation of quinic acid during biosynthesis of shikimic acid results from a microbe-catalyzed equilibration involving transport of initially synthesized shikimic acid back into the cytoplasm and operation of the common pathway of aromatic amino acid biosynthesis in the reverse of its normal biosynthetic direction. Shikimic acid transport has been shown to subject to catabolite repression.[36a] Hence increasing glucose availability in the culture (glucose-rich conditions) can repress shikimic acid transport and minimize formation of quinic acid. Under these culture conditions, all constructs (Entries 1 and 2, Table 9) entered the stationary phase of their growth 18 h after inoculation of the fermentor.

In addition to shikimic acid, the concentrations of DHS and quinic acid were determined. The total yield of hydroaromatic products was based on the combined yields of shikimic acid, DHS, and quinic acid. Shikimic acid, DHS, and quinic acid were determined by $^1$H NMR using response factors determined with authentic samples of all products. *E. coli* SP1.1/pKD15.071B and KL3*aroKaroL*/pKD15.071B provided the comparative strains for evaluating the impact on the concentrations and yields of synthesized shikimic pathway products attendant with altering the progenitor strain.

Both *E. coli* SP1.1/pKD15.071B and KL3*aroKaroL*/pKD15.071B synthesized shikimate pathway products in 37% total yield (Table 9). Varying the progenitor strain had no impact on the yield of the synthesized shikimate pathway products. Based on this result, the difference between the achieved total yield of 51% for DHS and shikimate

pathway byproducts synthesized by *E. coli* KL3/pJY1.216A and yield of 37% for shikimic acid and shikimate pathway byproducts synthesized by *E. coli* KL3*aroKaroL*/pKD15.071B will focus on the effect of shikimate dehydrogenase.

**Table 9. Comparison of the yields and concentrations of synthesized shikimic acid and shikimate pathway byproducts as a function of the progenitor host under glucose-rich conditions.**

| Entry | Strain | SA[a] | DHS[a] | QA[a] | Total Yield[b] |
|---|---|---|---|---|---|
| 1 | SP1.1/pKD15.071B | 67 | 18 | 4.4 | 37% |
| 2 | KL3*aroKaroL*/pKD15.071B | 60 | 18 | 5.4 | 37% |

[a] Abbreviations: DHS (DHS), quinic acid (QA), shikimic acid (SA); Unit: g/L.
[b] (mol DHS + mol QA + mol SA)/(mol glucose consumed).

Shikimate dehydrogenase is inhibited by shikimic acid exhibiting linear mixed-type inhibition with an inhibition constant of 0.16 mM and may limit the yield of synthesized shikimate pathway products.[37] Alternatively, because shikimate dehydrogenase needs NADPH to act as a cofactor, the demand of ample NADPH inside the cells with the overexpression of shikimate dehydrogenase may also limit the yield of synthesized shikimate pathway products. To examine the first hypothesis, isolation of a shikimate dehydrogenase isozyme insensitive to shikimic acid was essential. To examine the second hypothesis, approaches to improve the NADPH supply inside the cells are needed.

In the literature, it is a proven concept that in *Corynebacterium glutamicum* the production of L-lysine correlates with increased intracellular NADPH generation.[38] NADPH can be generated by glucose 6-phosphate dehydrogenase in the pentose pathway, isocitrate dehydrogenase in the TCA cycle and malic enzyme. It has been proved that pentose pathway was the most important source for intracellular NADPH generation in *C. glutamicum* because the contribution of both malic enzyme[39] and

isocitrate dehydrogenase[38b, 39a] to the generation of NADPH is of minor importance. One metabolic engineering strategy to increase NADPH availability entails redirection of pathway by inactivation of *pgi*-encoded phosphoglucose isomerase, which is the first enzyme specific for glycolysis. When an *E. coli pgi* mutant was derived from an *E. coli* tryptophan production strain, the efficiency of tryptophan formation was doubled.[40] Applying the same strategy for *C. glutamicum* strain DSM5715 the L-lysine formation was increased by 1.7-fold.[41] It was also shown that the reducing power imbalance caused by NADPH overproduction in an *E. coli pgi* mutant could be recovered to some extent by introducing an NADPH-consuming pathway such as the poly-3-hydroxybutyrate biosynthetic pathway of *E. coli pgi*.[42] Thus it would be interesting to introduce a *pgi* mutant into the shikimate-synthesizing *E. coli* constructs to evaluate the impact of NADPH availability.

## Effect of pH on the production of DHS

### Introduction

In nature, a wide range of proton concentrations is encountered: from pH 1 in acidic sulfur springs to pH 11 in soda lakes. Bacteria exist in all of these environments, although the number of species that can grow in extremely acidic habitats (acidophilic bacteria) and those (alkalophilic bacteria) that can grow in extremely alkaline ones are more restricted than those (neutrophilic bacteria) that can grow over the mid range from approximately pH 5.0 to 9.0. *E. coli*, which is representative of the large neutrophilic

class, grows at a maximum rate between pH 6.0 and pH 8.0 and more slowly at half a pH unit or so beyond these limits.[43]

The pH is an important parameter of the rate of many types of reactions, not only ionization of acids and bases but also solvolysis and oxidoreduction. As a result, enzymes and other macromolecules function optimally only over a narrow range of pH, usually close to neutrality, a range that is much more restricted than the pH range of the external environment over which growth of the bacterium is possible.

*E. coli,* as well as many other neutrophiles, has evolved remarkably effective mechanisms for homeostasis of intracellular pH.[44] *E. coli* regulates its pH at 7.4 to 7.8 during growth over the external pH range of 5.0 to 9.0.[45] The transmembrane pH difference ($\Delta$pH) is a component of the proton potential, which drives processes of transport, motility, and coupling of respiration.[44] Most of these processes are "secondary", in the sense that they entail no covalent bond exchanges, but a few are "primary" because they are used to drive a chemical reaction. Conceptually speaking, the simplest are the various porters, membrane transport proteins that couple the downhill flux of protons into the cell to the uphill flux of some metabolite inward or outward.

There are three kinds of porters: uniporter, symport and antiporter. Uniport proteins transport a substrate down its electrochemical gradient and are unlinked to any ion. The glycerol uniporter of *E. coli* is one example, but there are not many other examples. Symport protein is exemplified by the galactose permease (GalP).[18] This porter protein mediates the movement of two substrates in the same direction, a molecule of sugar plus one proton. The obligatory coupling between the two fluxes ensures that the electrochemical driving force upon the proton powers the concurrent accumulation of

the sugar. Another case is the accumulation of $K^+$ ions; this is believed to occur by symport with $H^+$. Antiport protein is so articulated that the flux of protons into the cell powers the efflux of some substrate from the cytoplasm.

In this study, the pH was lowered from 7.0 down to 6.5 and to 6.0. The increased transmembrane pH difference (ΔpH) will increase the proton potential across the membrane. Its effect on the production of DHS was examined by *E. coli* strains using either galactose permease (*E. coli* JY1.3/pKL5.17A) or wild-type PTS (*E. coli* KL3/pJY1.216A) to transport glucose inside the cells. Under our fed-batch fermentation conditions, culture pH is easily controlled.

**Table 10. Yields and concentrations of synthesized DHS and shikimate pathway byproducts as a function of pH under glucose-rich conditions.**

| Entry | Strain | Host[a]/Plasmid[b] | pH[c] | DHS[d] (g/L) Yield[e] | DAH[d], DHQ[d], GA[d] (g/L) | Total Yield[f] | Glucose Uptake[g] (mmol) |
|---|---|---|---|---|---|---|---|
| 1 | JY1.3/pKL5.17A | *pts⁻galP⁺* | 7.0 | (60) 36% | (0.0) (5.7) (7.1) | 43% | 1222 |
| 2 | JY1.3/pKL5.17A | *pts⁻galP⁺* | 6.5 | (70) 37% | (0.0) (6.8) (3.1) | 43% | 1369 |
| 3 | JY1.3/pKL5.17A | *pts⁻galP⁺* | 6.0 | (77) 42% | (0.0) (6.8) (2.1) | 48% | 1242 |
| 4 | KL3/pJY1.216A | *pts⁺/ ppsA* | 7.0 | (69) 35% | (10.0) (12.0) (13.0) | 51% | 1372 |
| 5 | KL3/pJY1.216A | *pts⁺/ ppsA* | 6.0 | (83) 39% | (12.0) (14.1) (1.6) | 51% | 1623 |

[a] in addition to *aroE*353 *serA*::*aroB*. [b] in addition to $aroF^{FBR}$, $P_{aroF}$, *tktA*, and *serA* plasmid inserts. [c] pH of culture medium. [d] Abbreviations: DHS (DHS), 3-deoxy-D-*arabino*-heptulosonic acid (DAH), 3-dehydroquinic acid (DHQ), gallic acid (GA), acetic acid (AcOH). [e] (mol DHS)/(mol glucose consumed). [f] (mol DHS + mol DAH + mol DHQ + mol GA)/(mol glucose consumed). [g] glucose consumed throughout the fermentation.

### DHS titers and yields as a function of pH

The DHS-synthesizing strains were initially cultured at 1 L scale in a 2.0-L working volume fermentor for 42 h (KL3/pJY1.216A) or 60-72 h (JY1.3/pKL5.17A) at 36 °C. Glucose-rich culture conditions were employed where the concentration of glucose in the medium was maintained in the range of 55-170 mM throughout the run.

Dissolved oxygen was maintained at 20%. All the fermentations were run in duplicate and reported results represent an average of two runs.

Substantial increases in the titer of DHS were observed (Table 10) with decreasing pH for both *E. coli* JY1.3/pKL5.17A and KL3/pJY1.216A cultured under glucose-rich conditions. With the change of pH from 7 to 6, approximately 15 g/L more DHS was synthesized for both strains (Entry 1 versus 3, Table 10; Entry 4 versus 5, Table 10). In addition, with the change of pH from 7 to 6, the total yield of the shikimate pathway products increased 5% for *E. coli* JY1.3/pKL5.17A (Entry 1 versus 3, Table 10) while remaining the same for *E. coli* KL3/pJY1.216A (Entry 4 versus 5, Table 10). The glucose consumption for *E. coli* JY1.3/pKL5.17A didn't change much at low pH while the glucose consumption for *E. coli* KL3/pJY1.216A increased about 20% at low pH (data not shown). The relationship between pH and glucose consumption by wild-type PTS or galactose permease is unexpected. At low pH, the increased transmembrane pH difference ($\Delta$pH) will increase the proton potential across the membrane. Glucose transport mediated by galactose permease required proton potential and glucose consumption throughout the fermentation was expected to increase with the increased proton potential at low pH. Glucose transport mediated by wild-type PTS didn't require proton potential to transport glucose and glucose consumption throughout the fermentation and was expected to stay the same at low pH.

Another interesting observation is the decrease of gallic acid accumulation at low pH. Gallic acid was accumulated normally during the stationary phase of the fermentation. At pH 6.0, both *E. coli* JY1.3/pKL5.17A and KL3/pJY1.216A accumulated only 2 g/L gallic acid compared to more than 7 g/L gallic acid at pH 7.0.

Gallic acid formation was due to either enzymatic or chemical catalysis.[46] In vitro DHS can react with $O_2$ in a phosphate buffer to form gallic acid, protocatechuic acid, tricarballyic acid and pyrogallol. Because the reaction of DHS was a general base-catalyzed process first order in both DHS and inorganic phosphate with an overall second order rate constant of $1.4 \times 10^{-5}$ $M^{-1}$ $s^{-1}$,[46] the acid environment at pH 6.0 may prevent DHS from being converted to gallic acid and other byproducts. This is probably the major reason for increased accumulation of DHS at low pH.

In conclusion, *E. coli* KL3/pJY1.216A and JY1.3/pKL5.17A synthesized higher concentrations of DHS at pH 6.0 relative to at pH 7.0. The 83 g/L of DHS produced by *E. coli* KL3/pJY1.216A at pH 6 is by far the highest titer achieved for DHS production because of both 20% increase in glucose consumption and decrease in gallic acid accumulation. However, the reason why glucose transport mediated by wild-type PTS is more efficient at low pH and whether lowering the pH can become a general strategy to increase the titer of shikimate pathway product remains to be investigated.

## REFERENCE


1 Niu, W.; Draths, K. M.; Frost, J. W. Benzene-Free Synthesis of Adipic Acid. *Biotechnol. Prog.* **2002**, *18*, 201-211.

2 Gibson, J. M.; Thomas, P. S.; Thomas, J. D.; Barker, J. L.; Chandran, S. S.; Harrup, M. K.; Draths, K. M.; Frost, J. W. Benzene-Free Synthesis of Phenol. *Angew. Chem., Int. Ed.* **2001**, *40*, 1945-1948.

3 Draths, K. M.; Frost, J. W. Environmentally Compatible Synthesis of Catechol from D-Glucose. *J. Am. Chem. Soc.* **1995**, *117*, 2395-2400.

4 Ran, N.; Knop, D. R.; Draths, K. M.; Frost, J. W. Benzene-Free Synthesis of Hydroquinone. *J. Am. Chem. Soc.* **2001**, *123*, 10927-10934.

5 (a) Barker, J. L.; Frost, J. W. Microbial Synthesis of *p*-Hydroxybenzoic Acid from Glucose. *Biotechnol. Bioeng.* **2001**, *76*, 376-390. (b) Amaratunga, M.; Lobos, J. H.; Johnson, B. F.; Williams, D. Genetically Engineered Microorganisms and Method for Producing 4-Hydroxybenzoic Acid. **2000** US Patent 6030819.

6 Li, K.; Frost, J.W. Synthesis of Vanillin from Glucose. *J. Am. Chem. Soc.* **1998**, *120*, 10545-10546.

7 Ensley, B. D.; Ratzkin, B. J.; Osslund, T. D.; Simon, M. J.; Wackett, L. P.; Gibson, D. T. Expression of Naphthalene Oxidation Genes in *Escherichia coli* Results in the Biosynthesis of Indigo. *Science*, **1983**, *222*, 167-169.

8 (a) Kambourakis, S.; Draths, K. M.; Frost, J. W. Synthesis of Gallic Acid and Pyrogallol from Glucose: Replacing Natural Product Isolation with Microbial Catalysis. *J. Am. Chem. Soc.* **2000**, *122*, 9042-9043. (b) Kambourakis, S.; Frost, J. W. Synthesis of Gallic Acid: $Cu^{2+}$-Mediated Oxidation of DHS. *J. Org. Chem.* **2000**, *65*, 6904-6909.

9 Chang, Y.; Almy, E. A.; Blamer, G. A.; Gray, J. I.; Frost, J. W.; Strasburg, G. M. Antioxidant Activity of DHS in Liposomes, Emulsions, and Bulk Oil. *J. Agric. Food Chem.* **2003**, *51*, 2753-2757.

10 (a) Ogino, T.; Garner, C.; Markley, J. L.; Herrmann, K. M. Biosynthesis of aromatic compounds: $^{13}C$ NMR spectroscopy of whole *Escherichia coli* cells. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5828-5832. (b) Weaver, L. M.; Herrmann, K. M. Cloning of an AroF Allele Encoding a Tyrosine-Insensitive 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase. *J. Bacteriol.***1990**, *172*, 6581-6584.

11 Li, K.; Mikola, M. R.; Draths, K. M.; Worden, R. M.; Frost, J. W. Fed-Batch Fermentor Synthesis of DHS Using Recombinant *Escherichia coli*. *Biotechnol. Bioeng.* **1999**, *64*, 61-73.

12 (a) Draths, K. M.; Frost, J. W. Synthesis Using Plasmid-Based Biocatalysis: Plasmid Assembly and 3-deoxy-D-*arabino*-heptulosonate Production. *J. Am. Chem. Soc.* **1990**, *112*, 1657-1659. (b) Draths, K. M.; Pompliano, D. L.; Conley, D. L.; Frost, J. W.; Berry, A.; Disbrow, G. L.; Staversky, R. J.; Lievense, J. Biocatalytic Synthesis of Aromatics from D-Glucose: The Role of Transketolase. *J. Am. Chem. Soc.* **1992**, *114*, 3956-3962.

13 (a) Li, K.; Frost, J. W. Microbial Synthesis of DHS: A Comparative Analysis of D-Xylose, L-Arabinose, and D-Glucose carbon sources. *Biotechnol. Prog.* **1999**, *15*, 876-883. (b) Li, K.; Frost, J. W. Utilizing Succinic Acid as a Glucose Adjunct in Fed-Batch Fermentation: Is Butane a Feedstock Option in Microbe-Catalyzed Synthesis? *J. Am. Chem. Soc.* **1999**, *121*, 9461-9462.

14 Postma, P. W.; Lengeler, J. W.; Jacobson, G. R. Phosphoenolpyruvate:Carbohydrate Phosphotranferase Systems. In *Escherichia coli and Salmonella: Cellular and Molecular Biology, 2nd Ed.*; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; pp 1149-1174.

15 (a) Patnaik, R.; Liao, J. C. Engineering of *Escherichia coli* Central Metabolism for Aromatic Metabolite Production with Near Theoretical Yield. *Appl. Environ. Microbiol.* **1994,** *60*, 3903-3908. (b) Liao, J. C.; Hou, S. Y.; Chao, Y. P. Pathway Analysis, Engineering, and Physiological Considerations for Redirecting Central Metabolism. *Biotechnol. Bioeng.* **1996**, *52*, 129-140.

16 Niersbach, M.; Kreuzaler, F.; Geerse, R. H.; Postma, P. W.; Hirsch, H. J. Cloning and Nucleotide Sequence of the *Escherichia coli* K-12 *PpsA* Gene Encoding PEP Synthase. *Mol. Gen. Genet.* **1992**, *231*, 332-336.

17 Parker, C.; Barnell, W. O.; Snoep, J. L.; Ingram, L. O.; Conway, T. Characterization of the *Zymomonas mobilis* glucose facilitator gene product (*glf*) in recombinant *Escherichia coli*: Examination of transport mechanism, kinetics and the role of glucokinase in glucose transport. *Mol. Microbiol.* **1995**, *15*, 795-802.

18 (a) Flores, N.; Xiao, J.; Berry, A.; Bolivar, F.; Valle, F. Pathway Engineering for the Production of Aromatic Compounds in *Escherichia coli*. *Nature Biotechnol.* **1996**, *14*, 620-623. (b) Saier, M. H. Jr.; Bromberg, F. G.; Roseman, S. Characterization of constitutive galactose permease mutants in *Salmonella typhimurium*. *J. Bacteriol.* **1973**, *113*, 512-514.

19 Pao, S. S.; Paulsen, I. T.; Saier, M. H. Jr. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 1-34.

20 Snell, K. D.; Draths, K. M.; Frost, J. W. Synthetic Modifica
*Escherichia coli* Chromosome: Enhancing the Biocatalytic Conversion of
Aromatic Chemicals. *J. Am. Chem. Soc.* **1996**, *118*, 5605-5614.

21 Cobbett, C. S.; Delbridge, M. L. Regulatory Mutants of the AroF-Tyr
*Escherichia coli* K-12. *J. Bacteriol.* **1987**, *169*, 2500-2506.

22 (a) Konstantinov, K. B.; Nishio, N.; Yoshida, T. Glucose Feedi
Accounting for the Decreasing Oxidative Capacity of Recombinant *Escher*
Fed-Batch Cultivation for Phenylalanine Production. *J. Ferment. Bioeng.* **19**
260. (b) Konstantinov, K. B.; Nishio, N.; Seki, T.; Yoshida, T. Physiologicall
Strategies for Control of the Fed-Batch Cultivation of Recombinant *Escheri*
Phenylalanine Production. *J. Ferment. Bioeng.* **1991**, *71*, 350-355. (c) Kle
Strohl, W. R. Acetate Metabolism by *Escherichia coli* in High-C
Fermentation. *Appl. Environ. Microbiol.* **1994**, *60*, 3952-3958.

23 Meyer, D.; Schneider-Fresenius, C.; Horlacher, R.; Peist, R.; Boos, W
Characterization of Glucokinase from *Escherichia coli* K-12. *J. Bacteriol*
1298-1306.

24 (a) Postma, P. W.; Lengeler, J. W.; Jacobson, G. R. Phosphoen
Carbohydrate Phosphotransferase Systems. In *Escherichia coli and Salmone*
*and Molecular Biology*, 2nd ed.; Neidhardt, F. C., Ed.; ASM Press: Wash
1996; pp 1149-1174. (b) Cordaro, J. C.; Melton, T.; Stratis, J. P.; Atagün, M
C.; Hartman, P. E.; Roseman, S. Fosfomycin Resistance: Selection Method
and Extended Deletions of the Phosphoenolpyruvate:Sugar Phosphotransfera
*Salmonella typhimurium. J. Bacteriol.* **1976**, *128*, 785-793.

25 Flores, N.; Xiao, J.; Berry, A.; Bolivar, F.; Valle, F. Pathway Enginee
Production of Aromatic Compounds in *Escherichia coli. Nature Biotechno*
620-623.

26 (a) Knop, D. R.; Draths, K. M.; Chandran, S. S.; Barker, J. L.; von D
Weber, W.; Frost, J. W. Hydroaromatic Equilibration During Biosynthesis
Acid. *J. Am. Chem. Soc.* **2001**, *123*, 10173-10182. (b) Draths, K. M.; Knop,
J. W. Shikimic Acid and Quinic Acid: Replacing Isolation from Plant S
Recombinant Microbial Biocatalysis. *J. Am. Chem. Soc.* **1999**, *121*, 1603-160

27 Patnaik, R.; Roof, W. D.; Young, R. F.; Liao, J. C. Stimulation
Catabolism in *Escherichia coli* by a Potential Futile Cycle. *J. Bacteriol.* **1992**
7532.

28 El-Mansi, E. M. T.; Holms, W. H. Control of carbon flux to aceta
during growth of *E. coli* in batch and continuous culture. *J. Gen. Microbiol*
2875-2883.

29 Farmer, W. R.; Liao, J. C. Reduction of Aerobic Acetate Production by *Escherichia coli*. *Appl. Environ. Microbiol.* **1997**, *63*, 3205-3210.

30 (a) Hogema, B. M.; Arents, J. C.; Bader, R.; Eijkemans, K.; Inada, T.; Aiba, H.; Postma, P. W. Inducer exclusion by glucose 6-phosphate in *Escherichia coli*. *Mol. Microbiol.* **1998**, *28*, 755-765.

31 Hernandez-Montalvo, V.; Valle, F.; Bolivar, F.; Gosset, G. Characterization of sugar mixtures by an *Escherichia coli* mutant devoid of the phosphotransferase system. *Appl. Microbiol. Biotechnol.* **2001**, *57*, 186-191.

32 (a) Backman, K.C. U.S. Patent 5,169,768, 1992. (b) Miller, J. E.; Backman, K.C.; O'Conner, M. J.; Hatch, R. T. Production of Phenylalanine and Organic-Acids by Phosphoenolpyruvate Carboxylase-Deficient Mutants of *Escherichia coli*. *J. Ind. Microbiol.* **1987,** 2, 143-149.

33 Mori, M.; Yokota, A. ; Sugitomo, S.; Kawamura, K. Patent JP 62,205,782, 1987.

34 Yi, J.; Li, K.; Draths, K. M.; Frost, J. W. Modulation of Phosphoenolpyruvate Synthase Expression Increases Shikimate Pathway Product Yields in *E. coli*. *Biotechnol. Prog.* **2002**, *18*, 1141-1148.

35 Chandran, S. S.; Yi, J.; Draths, K. M.; von Daeniken, R.; Weber, W.; Frost, J. W. Phosphoenolpyruvic Acid Availability and the Biosynthesis of Shikimic Acid. *Biotechnol. Prog.* **2003**, *19*, 808-814.

36 (a) Knop, D. R.; Draths, K. M.; Chandran, S. S.; Barker, J. L.; von Daeniken, R.; Weber, W.; Frost, J. W. Hydroaromatic Equilibration During Biosynthesis of Shikimic Acid. *J. Am. Chem. Soc.* **2001**, *123*, 10173-10182. (b) Draths, K. M.; Knop, D. R.; Frost, J. W. Shikimic Acid and Quinic Acid: Replacing Isolation from Plant Sources with Recombinant Microbial Biocatalysis. *J. Am. Chem. Soc.* **1999**, *121*, 1603-1604.

37 Dell, K. A.; Frost, J. W. Identification and removal of impediments to biocatalytic synthesis of aromatics from D-glucose: Rate-limiting enzymes in the common pathway of aromatic amino acid biosynthesis. *J. Am. Chem. Soc.* **1993**, *115*, 11581-11589.

38 (a) Genealogy profiling through strain improvement by using metabolic network analysis: metabolic flux genealogy of several generations of lysine-producing corynebacteria. *Appl. Environ. Microbiol.* **2002**, *68*, 5843-5849. (b) Marx, A.; de Graaf, A. A.; Wiechert, W.; Eggeling, L.; Sahm, H. Metabolic fluxes in *Corynebacterium glutamicum*-identification of metabolic patterns by $^{13}C$ isotope analysis. In: Preprints of the Seventh International Conference on Computer Applications in Biotechnology. May 31 to June 4, 1998, Osaka, Japan, pp. 387-392. (c) Sonntag, K.; Schwinde, J.; deGraaf, A. A.; Marx, A.; Eikmanns, B. J.; Wiechert, W.; Sahm, H. $^{13}C$ nuclear magnetic resonance

studies of the fluxes in the central metabolism of *Corynebacterium glutamicum* during growth and overproduction of amino acids in batch cultures. *Appl. Microbiol. Biotechnol.* **1995**, *44*, 489-495.

39 (a) Marx, A.; Striegel, K.; de Graaf, A. A.; Sahm, H.; Eggeling, L. Response of the central metabolism of *Corynebacterium glutamicum* to different flux burdens. *Biotechnol. Bioeng.* **1997**, *56*, 168-180. (b) Marx, A.; Eikmanns, B. J.; Sahm, H.; de Graaf, A. A.; Eggeling, L. Response of the central metabolism in *Corynebacterium glutamicum* to the use of an NADH-dependent glutamate dehydrogenase. *Metabol. Eng.* **1999**, *1*, 35-48.

40 Deletion of *pgi* alters tryptophan biosynthesis in a genetically engineered strain of *Escherichia coli*. *Appl. Environ. Microbiol.* **1991**, *57*, 2995-2999.

41 (a) Dunican, L. K.; Mccormback, A.; Stapelton, C.; Burke, K.; O'Donohue, M.; Marx, A.; Mockel, B. Cloning and uses of a novel nucleotide sequence coding for glucose-6-phosphate isomerase (*pgi*) from bacteria. 2001, Patent EP 1087015. (b) Marx, A.; Hans, S.; Mockel, B.; Bathe, B.; deGraaf, A. A. Metabolic phenotype of phosphoglucose isomerase mutants of *Corynebacterium glutamicum*. *J. Biotechnol.* **2003**, *104*, 185-197.

42 (a) Shi, H.; Nikawa, J.; Shimizu, K. Effect of modifying metabolic network on ply-3-hydroxybutyrate biosynthesis in recombinant *Escherichia coli*. *J. Biosci. Bioeng.* **1999**, *87*, 666-677. (b) Kabir, M. M.; Shimizu, K. Gene expression patterns for metabolic pathway in *pgi* knockout *Eshcerichia coli* with and without *phb* genes based on RT-PCR. *J. Biotechnol.* **2003**, *105*, 11-31.

43 Ingraham, J. L.; Marr, A. G. Effect of Temperature, Pressure, pH, and Osmotic Stress on Growth. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd ed.; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; pp 1570-1578.

44 (a) Booth, I. R. Regulation of cytoplasmic pH in bacteria. *Microbiol. Rev.* **1985**, *49*, 359-378. (b) Padan, E.; Schuldiner, S. Intracelluar pH and membrane potential as regulators in the prokaryotic cell. *J. Membr. Biol.* **1987**, *95*, 189-198.

45 (a) Padan, E.; Zilberstein, D.; Rottenberg, H. The proton electrochemical gradient in *Escherichia coli* cells. *Eur. J. Biochem.* **1976**, *63*, 533-541. (b) Slonczewski, J. L.; Rosen, B. P.; Alger, J. R.; macnab, R. M. pH homeostasis in *Escherichia coli*: measurement by $^{31}P$ nuclear magnetic resonacnce of methylphosphonate and phosphate. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 6271-6275. (c) Zilberstein, D.; Agmon, V.; Schuldiner, S.; Padan, E. *Escherichia coli* intracellular pH, membrane potential, and cell growth. *J. Bacteriol.* **1984**, *158*, 246-252.

46 Kambourakis, S. Synthesis of Value-Added Chemicals from Glucose Using Chemical and Microbial Catalysis: Gallic acid, Protocatechuic acid, Pyrogallol and Catechol. Ph. D. thesis. Michigan State University, East Lansing, MI. 2000.

# CHAPTER 3

## Microarray analysis of gene expression changes in response to overexpression of PEP synthase

### Introduction

Researchers have employed a variety of strategies for improving the yield of shikimate pathway product in *E. coli*. The availability of PEP acid emerges as a key factor limiting carbon flow directed into the shikimate pathway after overexpression of *aroF*$^{FBR}$-encoded DAHP synthase and *tktA*-encoded transketolase to remove feedback inhibition and ameliorate restricted availability of D-erythrose 4-phosphate.[1] One strategy to surmount the limitation of PEP employed overexpression of *E. coli ppsA*-encoded PEP synthase to recycle PTS-generated pyruvic acid back to phosphoenolpyruvic acid, as was discussed in the preceding chapter.[1a] The resulting strain *E. coli* KL3/pJY1.216A is currently synthesizing the highest titer and yield of biosynthesis of 3-dehydroshikimic acid when PEP synthase is expressed at optimal level. However, the achieved total yield 51% (mol/mol) of 3-dehydroshikimic acid and shikimate pathway byproducts synthesized from glucose was still far removed from the theoretical maximum 86% (mol/mol).

In order to make a better biocatalyst to produce more 3-dehydroshikimic acid, expression arrays was used as a tool to investigate both the consequences of PEP synthase overexpression during the time course of the fermentation and the expression profile changes between different stages of the fermentation. The comparison between "relative high yield" cells when IPTG was added to induce the overexpression of PEP

synthase and "relative low yield" cells when no IPTG was added may provide clues to limitations in metabolic flux. In addition, the comparison of expression profiles between different stages of the fermentations will help us understand the regulation of the genes during the time course of the fermentation and may provide clues to limitations in metabolic flux particularly during the end of the fermentation. An advantage for this analysis is that the same strain *E. coli* KL3/pJY1.216A was used to generate all the expression array data thus prevent the complications from different strain background. This method is limited, however, by an inability to detect mutations that alter primary sequence (and function) without affecting mRNA levels. For example, *aroF* and *aroF*$^{FBR}$ cannot be differentiated by this method. Where examined, previous studies have indicated that expression changes provide a good predictor for metabolic flux[2] and changes in protein levels.[3] Recent improvements in methodology and simple statistical tools now allow the reproducible measurement of changes that differ by less than 2-fold.[2a, 4]

**Fed-batch fermentation of *E. coli* KL3/pJY1.216A for RNA isolation**

*E. coli* KL3/pJY1.216A, which was discussed in detail in Chapter 2, was cultured under glucose-rich conditions at 36°C and pH 7.0 in a 2.0-L working volume fermentor. The concentration of glucose in the medium was maintained in the range of 55-170 mM throughout the run. Dissolved oxygen was maintained at 20% air saturation by allowing the impeller speed to vary. To induce PEP synthase expression, 12 mg IPTG was added when the preset maxima of 750 rpm (stir rate) and 1.0 L/L/min (air flow rate) were

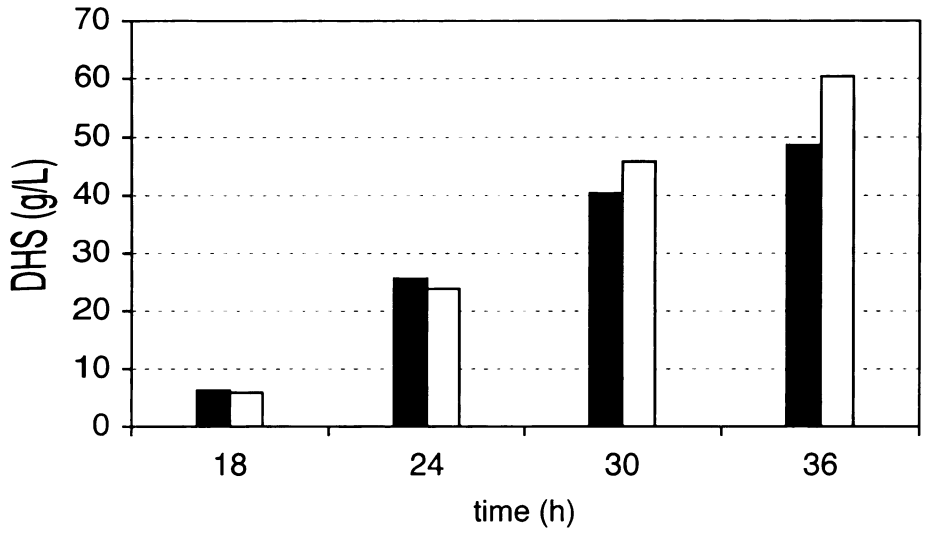



**Figure 23. DHS production by *E. coli* KL3/pJY1.216A with and without IPTG addition.** Abbreviation: 3-dehydroshikimic acid (DHS). No IPTG added (black bars), 12 mg of IPTG added at 16 h, 22 h, 28 h and 30 h (open bars).

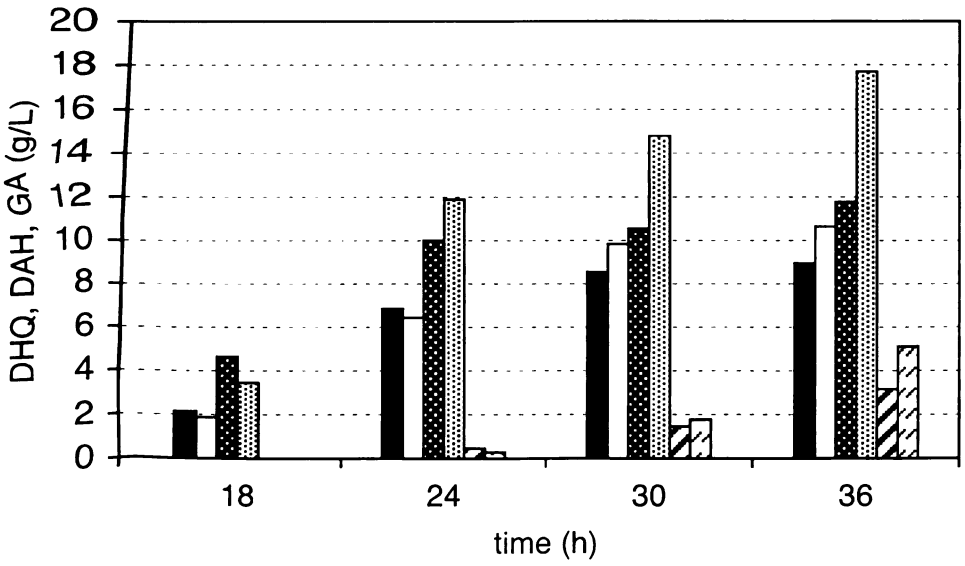



**Figure 24. DHQ, DAH and GA production by *E. coli* KL3/pJY1.216A with and without IPTG.** Abbreviations: 3-deoxy-D-*arabino*-heptulosonic acid (DAH), 3-dehydroquinic acid (DHQ), gallic acid (GA). No IPTG added, DHQ (black bars); IPTG added, DHQ (open bars); No IPTG added, DAH (black bars with dots); IPTG added, DAH (open bars with dots); No IPTG added, GA (hashed bars); IPTG added, GA (open bars with curves).

reached (usually 16 h), and every 6 h after the first addition (usually 22 h, 28 h, 34 h). In 36 h, a total yield of 39% (mol/mol) of DHS (Figure 23) and shikimate pathway byproduct (Figure 24) was realized by *E. coli* KL3/pJY1.216A in the absence of IPTG addition. This contrasts with the total yield of 52% (mol/mol) when *E. coli* KL3/pJY1.216A was cultured with IPTG addition (Figure 23 and 24). Glucose-rich fed-batch fermentations were run in triplicate and reported results represent an average of three runs.

**Microarray platform**

Under the discovery-based approach, DNA microarrays are used as screening tools to identify genes associated with the biological process of interest. Once target genes are identified, additional experiments may be designed to validate this list. The efficiency of knowledge discovery using this high-throughput experimental process depends upon the reliability of the microarray technology used in the initial screening experiments. Research planning to utilize microarray experiments for discovery-based research must evaluate available commercial technologies when allocating laboratory resources for prospective experiments. As discussed in the introduction chapter, Panorama *E. coli* gene array was chosen to be the platform to perform microarray analysis.

Each Panorama array contains PCR-amplified open reading frames (ORFs) from the *E. coli* K-12 (MG1655) genome. The majority of ORFs have been amplified from start to stop codon. All of the 4,290 ORFs have been printed in duplicate at equal mass per spot onto positively charged nylon hybridization membranes. Following printing,

each array was cross-linked with UV light. The arrays are 12 cm × 24 cm in size. The genomic DNA is found in the corner of the arrays. During any expression profiling experiment where labeled cDNA is used as probe, the genomic DNA spots will act as a positive control and should always show significant signals. However, this signal strength was not utilized in the analysis in this study.

**Synthesis of Radiolabeled cDNA**

RNA was purified from the cells at 18h, 24 h, 30 h and 36 h after initiation of the run from the triplicate fermentation in which IPTG was added and the triplicate fermentation in which IPTG was not added. Total RNA was isolated using the hot phenol method and purified as described by Tao et al.[4a] Radioactive cDNA was synthesized as described by Gonzalez et al.[2a] using [α-$^{33}$P]dCTP (New England Nuclear) and a set of primers homologous to the 3' ends of the predicted 4,290 ORFs in *E. coli* (Sigma-Genosys). Approximately 70% of the added radioactivity was incorporated into cDNA.

**Hybridization**

Samples of cDNA were hybridized to DNA probes on nylon membranes (Sigma-Genosys). Hybridizations were performed in roller bottles at 65 °C. A total of eight filters used in these experiments were from the same manufacturing lot. Each filter was stripped and used 3 times. Each filter corresponds to each condition investigated (with and without IPTG addition, each at 4 time points during the fermentation for a total of 8 different conditions).

**Quantitation of cDNA as a relative measure of gene expression (mRNA)**

Wet membranes were wrapped in Mylar film and exposed to a phosphorimager screen (Molecular Dynamics) for 20 h. Exposed screens were scanned using a phosphoimager at a pixel size of 50 μm. Array Vision software (version 6.1, Imaging Research Inc.) was used to quantify the intensity of each spot after correcting automatically for background using the "surrounding entire template" method. The sum of values for all 4,290 ORFs in each array was normalized to 100 million to allow comparison of gene expression between experiments. Each ORF is expressed as a percentage of the total (i.e., 100,000 corrsponds to 0.1% of total cDNA hybridized to the array). Results for every time point represent an average of six determinations (three hybridizations each, two arrays per filter). Statistical analyses and data transformations were performed in Excel using the log of expression values.

**Statistical analysis**

Among the 4290 ORFs, 347 genes with an average expression level below 0.002% were considered unreliable. The coefficient of variation (standard deviation expressed as a percentage of the mean) was used to estimate the error associated with the measurement of the remaining ORFs. Table 11 lists the average coefficient of variation for each data set. Considerable error exists in these measurements, with an average coefficient of variation of 26%.

As pointed out by Arfin et al.,[5] the significance of gene expression comparisons cannot be interpreted based solely on the ratio of measured values without statistical evaluation. The same parametric statistical methods that are routinely used with small

sets of biological and biochemical data can be employed provided gene expression data can be shown to follow a normal distribution. Log expression has previously been shown to approximate a normal distribution.[2a] In our studies, log values of expression data were generated and statistical significance of gene comparisons at the same time point between "high yield" cells and "low yield" cells was evaluated by computing probability values for the null hypothesis using two-tailed student *t* tests.[6] This test produces a *p* value that represents the probability that the difference is observed because of random chance. A very small *p* value ($p < 0.05$) will indicate that the tested gene is likely to be differentially expressed. Using student *t* test, many smaller but significant changes were identified.

**Table 11. Comparison of data reproducibility (coefficient of variation) for all the samples obtained from *E. coli* KL3/pJY1.216A fermentation[a].**

| | 18 h | 24 h | 30 h | 36 h |
|---|---|---|---|---|
| No IPTG | 23% | 26% | 30% | 30% |
| With IPTG | 25% | 23% | 30% | 22% |

[a] Results for each condition are summarized as a coefficient of variation (CV), the standard deviation expressed as a percentage of the mean. Values in each column represent the average CV for all ORFs under condition. Genes with expression value below 0.002% of the total message (threshold for reliable measurement[2a]) were eliminated from consideration.

**Cluster analysis to define corresponding physiological states**

Clusting analysis[7] is the standard means by which large sets of experiments and transcripts are analyzed. Generally, some measure of similarity is used to place each gene (and/or experiment) into a single group or a position within a hierarchy. The output can be displayed as lists, but are more typically visualized by some variation of the

red/green matrix display originally introduced by Eisen *et al* .[8] An interesting application of this approach is the clustering of tumors to find new possible tumor subclasses.[9]

Cluster analysis was used to examine if there are significant global expression differences between samples of all conditions and identification of the most appropriate pairing of samples for comparison. Hierarchical clustering by average linkage[8] was carried out to intuitively examine gene expression variation among samples of different conditions using Cluster software downloaded from Standford University. Basically, this is an agglomerative process in which single-member clusters are fused to bigger and bigger clusters. In somewhat more detail, the procedure starts by computing a pairwise distance matrix between all the genes or conditions, the distance matrix is explored for the nearest genes or conditions, and they are defined as a cluster. There are several hierarchical clustering algorithms that differ in the way the distances are calculated. After a new cluster is formed by agglomeration of two clusters, the distance matrix is updated to reflect its distance from all other clusters. Then, the procedure searches for the nearest pair of clusters to agglomerate, and so on. This procedure leads to a hierarchical dendogram in which multiple clusters are fused in nodes according to their similarity, finally resulting in a single hierarchical tree (Figure 25). Relationships among objects (conditions) are represented by a tree whose branch lengths reflect the degree of similarity between the objects (conditions), as assessed by a pairwise similarity function.

In agreement with growth state, the samples could be clustered into 2 big groups (early fermentation at 18 h and 24 h; later stages of fermentation at 30 h and 36 h) and 4 small groups (Figure 25). For more convenient analysis of transcriptome profile variations, the entire cultivation was divided into four growth stages as follows: 1.

exponential growth phase in which cells grow at a constant specific growth rate (18 h); 2. early stationary phase (24 h); 3. middle stationary phase (30 h); 4. late stationary phase (36 h). It is interesting to notice that, at 30 h and 36 h, expression profiles for *E. coli* KL3/pJY1.216A with and without IPTG addition were not clustered together as expected. On the contrary, expression profiles for *E. coli* KL3/pJY1.216A with IPTG addition at 30 h and 36 h were clustered together. Expression profiles for *E. coli* KL3/pJY1.216A without IPTG addition at 30 h and 36 h were also clustered together. This result indicates that there are significant changes of physiological states at 30 h and 36 h with PEP synthase overexpression relative to the control. Based on cluster analysis, pairwise comparisons of the "high yield" cells and "low yield" cells are appropriate for the 18 h and 24 h samples. Comparisons at 30 h and 36 h samples appear valid but may be anomalous.

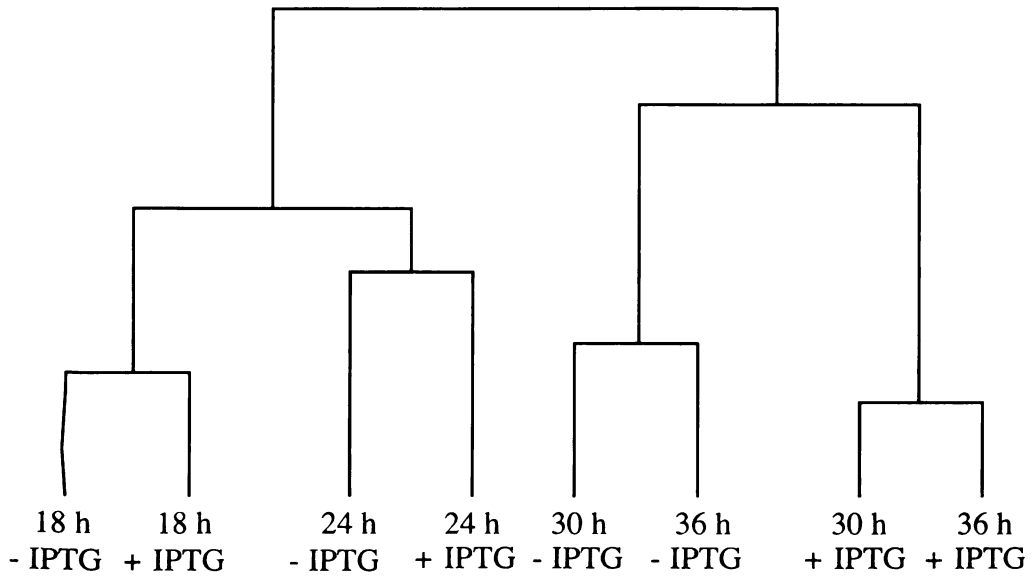



**Figure 25. Cluster analysis of expression data by each condition.**

**Table 12. Number of differentially expressed ORFs in response to PEP synthase (PpsA) overexpression.**

| Time points | 18 h | 24 h | 30 h | 36 h |
|---|---|---|---|---|
| The total genes with $p$-values < 0.05 | 716 | 1150 | 747 | 2186 |
| Increased by 2-fold or more with PpsA overexpression | 27 | 227 | 53 | 214 |
| Decreased by 2-fold or more with PpsA overexpression | 17 | 52 | 169 | 1056 |

**How many genes are differentially expressed with phosphenolpyruvate synthase overexpression during the fermentation?**

A simple method for finding sets of interesting genes that may provide clues to limitations in metabolic flux is comparing expression profiles of two or more samples for differentially expressed genes. Tables 12 through 14 summarize comparisons for gene array data. Expression results were compared for *E. coli* KL3/pJY1.216A during the fermentation of glucose either with or without IPTG induction of phosphenolpyruvate synthase overexpression under the glucose-rich conditions at individual time points. Comparisons were made on a statistical basis and are presented as a ratio. Table 12 indicated total ORFs that differed significantly (defined as $p < 0.05$) and total ORFs that differed by 2-fold or greater for each comparison (also $p < 0.05$). As a result of the relatively low standard deviations for ORFs in these data sets, the statistical approach identified many more differentially expressed ORFs than the ratio approach. Table 13 and Table 14 break down those ORFs with 2-fold or greater difference by functional groups based on the information provided by Sigma Genosys. The names of ORFs that differed by 2-fold or greater simultaneously at two individual time points (i.e. at both 24

h and 30 h) were also shown in Table 13 and Table 14. Identification of these ORFs by functional groups served as a guide for further investigation.

**Table 13. Functional group of the upregulated expressed ORFs (ratio >2.0) in response to PEP synthase overexpression.**

| Functional group[a] | Number of genes (ratio > 2.0) | | | | | Gene name[c] (ratio > 2.0) | | |
|---|---|---|---|---|---|---|---|---|
| | Total[b] | 18h | 24h | 30h | 36h | 18h & 24h | 24h & 30h | 30h & 36h |
| Carbon compound metabolism | 129 | 0 | 4 | 2 | 12 | | *poxB* | *poxB* |
| Regulatory protein | 178 | 1 | 7 | 2 | 28 | | *cyaA* | *b0705* |
| Phage, transposon or plasmid | 87 | 0 | 0 | 1 | 13 | | | |
| Structure protein | 236 | 1 | 5 | 2 | 22 | | | |
| Transport | 427 | 2 | 14 | 5 | 51 | | *crr msYB bfr* | *potH bfr* |
| Cell process | 197 | 0 | 22 | 4 | 20 | | *dps dnaK sodA mopB* | *dnaK sodA* |
| Energy metabolism | 243 | 0 | 19 | 8 | 25 | | *adhE* | |
| DNA replication, repair | 115 | 2 | 2 | 0 | 7 | | | |
| Transcription | 55 | 0 | 1 | 3 | 2 | | | *rpoB* |
| Translation | 182 | 1 | 1 | 3 | 8 | | | *rpiL* |
| Amino acid biosynthesis | 131 | 0 | 8 | 5 | 6 | | *trpB* | *trpA trpB trpC trpD trpE* |
| Nucleotide metabolism | 58 | 0 | 1 | 0 | 3 | | | |
| Fattic acid metabolism | 48 | 0 | 2 | 0 | 1 | | | |
| Central intermediary metabolism | 188 | 3 | 14 | 2 | 3 | *ppsA* | *gldA aceB* | *gldA aceB* |
| Cofactors and prosthetic groups | 103 | 2 | 5 | 0 | 3 | | | |
| Putative | 277 | 6 | 6 | 1 | 1 | *pgiA* | | |
| Hypothetical | 1635 | 9 | 116 | 15 | 9 | *ydfE ydiA b1836 hdeB hdeD hdeA yifE* | *b1005 b1112 b1783 yhbH b3238 b3242 yjdB* | *b1005 b1112 b3238 yjdB* |

[a] Functional Group was classified based on the information provided by Sigma Genosys.
[b] Total number of identified ORFs for this functional group.
[c] The names of gene that differed by 2-fold or greater simultaneously at two individual time points.

**Table 14. Functional group of the downregulated expressed ORFs (ratio >2.0) in response to PEP synthase overexpression.**

| Functional group[a] | Number of genes | | | | | Gene name | | |
|---|---|---|---|---|---|---|---|---|
| | Total[b] | 18h | 24h | 30h | 36h | 18h & 24h | 24h & 30h | 30h & 36h |
| Carbon compound metabolism | 129 | 1 | 2 | 1 | 8 | | | *ascG* |
| Regulatory protein | 178 | 0 | 5 | 1 | 12 | | | *atoS* |
| Phage, transposon or plasmid | 87 | 0 | 7 | 1 | 14 | | | |
| Structure protein | 236 | 1 | 7 | 1 | 25 | | | |
| Transport | 427 | 3 | 10 | 6 | 34 | | | *betT cydD gatC ugpA tnaB* |
| Cell process | 197 | 1 | 3 | 3 | 14 | | | *osmY rhiV* |
| Energy metabolism | 243 | 1 | 5 | 1 | 17 | | *hycC* | |
| DNA replication, repair | 115 | 0 | 2 | 4 | 21 | | | *modF priA* |
| Transcription | 55 | 0 | 0 | 3 | 16 | | | *sbcB sbcC hsdR* |
| Translation | 182 | 1 | 0 | 2 | 31 | | | *yihK* |
| Amino acid biosynthesis | 131 | 3 | 2 | 8 | 35 | | | *b1748 yagF* |
| Nucleotide metabolism | 58 | 4 | 1 | 7 | 16 | | *adk* | *apt adk purT* |
| Fattic acid metabolism | 48 | 0 | 0 | 4 | 12 | | | *fadA acs b4249* |
| Central intermediary metabolism | 188 | 0 | 1 | 7 | 41 | | | *pckA bisZ b0221 edd b2463 hisQ* |
| Cofactors and prosthetic groups | 103 | 1 | 0 | 3 | 16 | | | *menF hemY* |
| Putative | 277 | 0 | 1 | 18 | 85 | | | 16 genes |
| Hypothetic | 1635 | 1 | 4 | 99 | 659 | | *yebL* | 83 genes including *yebL* |

[a] Functional Group was classified based on the information provided by Sigma Genosys.
[b] Total number of identified ORFs for this functional group.
[c] The names of gene that differed by 2-fold or greater simultaneously at two individual time points.

## Pathway diagrams

There is substantial literature about molecular mechanisms and biochemical pathways,[10] and some information has been organized into metabolic pathway databases,

most notably EcoCyc,[11] but also KEGG,[12] and WIT.[13] Now it is possible to overlay all the expression data into all published pathway of *E. coli* metabolism in the EcoCyc

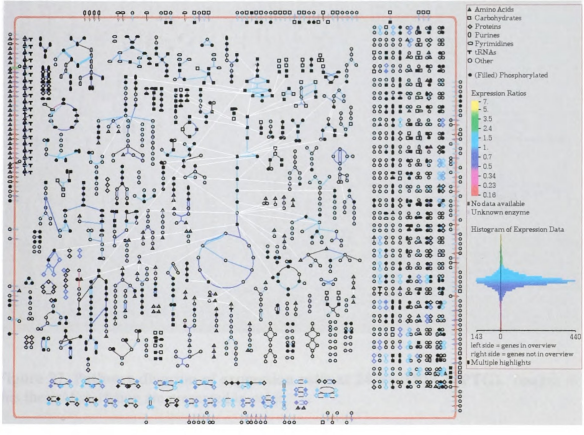



**Figure 26. Pathway diagram of expression ratio at 18 h (+IPTG/-IPTG).** Images in this thesis/dissertation are presented in color.

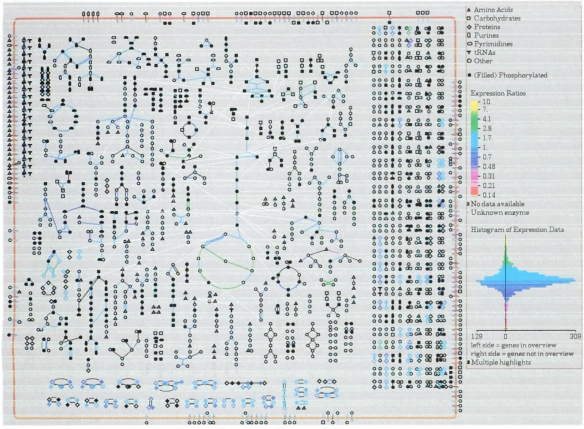



**Figure 27. Pathway diagram of expression ratio at 24 h (+IPTG/-IPTG).** Images in this thesis/dissertation are presented in color.

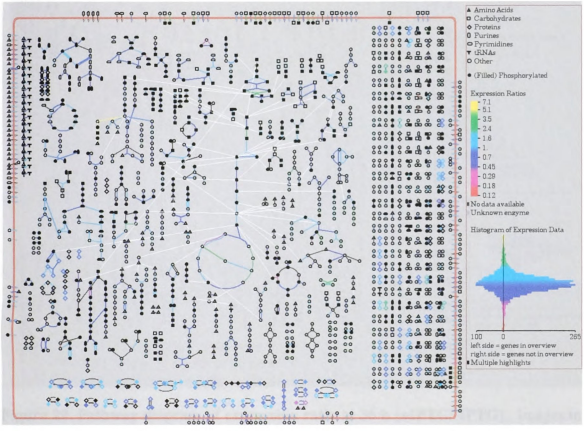



**Figure 28. Pathway diagram of expression ratio at 30 h (+IPTG/-IPTG).** Images in this thesis/dissertation are presented in color.

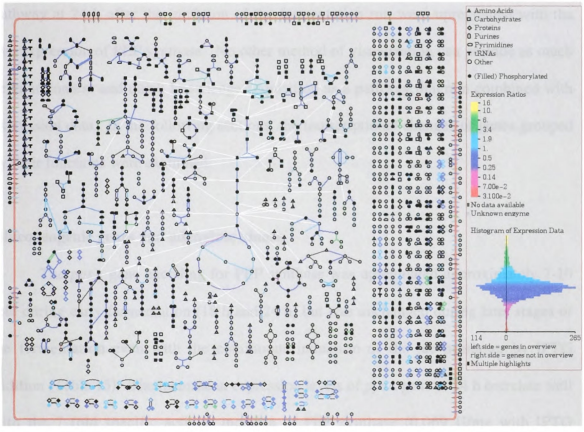



**Figure 29. Pathway diagram of expression ratio at 36 h (+IPTG/-IPTG).** Images in this thesis/dissertation are presented in color.

database. In this study at each time point, the expression ratio of two samples (PEP synthase overexpression and wild type PEP synthase expression) for each gene was overlaid onto all the pathways in the EcoCyc database (Figures 26, 27, 28 and 29). Reaction lines are color-coded according to the relative expression level of the gene that codes for the enzyme that catalyzes that reaction step. In the pathway diagram the blue, green and yellow color showed upregulation genes with PEP synthase overexpression. The red color showed downregulated genes with PEP synthase overexpression. Downloading expression ratios for each gene into the EcoCyc database provides a convenient visual analysis of expression data. For example, it can be seen instantly in the pathway diagram (Figure 27) that genes involved in the TCA cycle and glycolysis

pathway at 24 h after the initiation of the fermentation run were upregulated with the overexpression of PEP synthase. No other method of visualization can provide as much cellular context and quick knowledge acquisition as a pathway diagram combined with expression data. In the following sections, the transcriptional changes of genes grouped by their functions were examined.

**Gluconeogenic genes and adenylate kinase**

The *ppsA* gene encoding for PEP synthase was upregulated approximately 7-10 fold during early fermentation (18 h and 24 h) but was unchanged during later stages of the fermentation even with the continuous induction of the tac promoter by IPTG addition (Table 15). The increase in expression levels of *ppsA* gene at 24 h correlate well with the 7 fold specific activity increase of PEP synthase (0.069 U/mg with IPTG addition at 24 h, 0.0093 U/mg without IPTG addition at 24 h) as measured in Chapter 2 of this thesis. The increase in expression levels of *ppsA* gene at 30 h and 36 h without IPTG addition (from 0.04% of total cDNA at 24 h to approximately 0.15% of total cDNA at 30 h and 36 h) may reflect the cell's need for metabolic flux from pyruvate to PEP during the later stages of the fermentation. The underlying mechanism for this upregulation of *ppsA* gene at later stages of the fermentation without IPTG addition, however, remains to be determined. The *pckA* gene encoding PEP carboxykinase that converts the TCA cycle intermediate oxaloacetate to PEP was downregulated at all time points with the overexpression of PEP synthase. It is expected because there is ample supply of PEP with the overexpression of PEP synthase.

The *adk* gene (Table 15) encoding adenylate kinase was downregulated with the overexpression of PEP synthase at all time points. Combining pyruvate kinase, PEP synthase and adenylate kinase together catalyzes reactions in which ATP is converted to ADP and $P_i$ (Figure 30). When PEP synthase was overexpressed, the downregulation of *adk* can prevent the formation of such futile cycle that only consumes ATP.

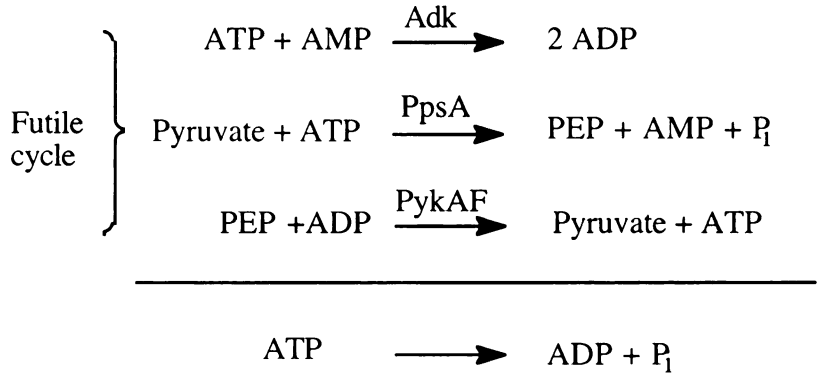



**Figure 30. Futile cycle involving Adk (adenylate kinase), PpsA (PEP synthase) and PykAF (pyruvate kinase).**

**Embden-Meyerhof pathway**

Table 15 summarizes the expression data for all 15 genes involved in the metabolism of glucose to pyruvate (Embden-Meyerhof pathway). Although results at each time point were normalized for total bound cDNA, higher values of expression level were observed for most of these genes during early fermentation (18 h and 24 h) than during the later stages of fermentation (30 h and 36 h). For many genes such as *pgi* encoding phosphoglucose isomerase, *fbaA* encoding fructose bisphoshate aldolase, *gapA* encoding glyceraldehydes 3-phosphate dehydrogenase, *pgk* encoding phosphoglycerate

**Table 15. Expression data for gluconeogenic genes and all the genes involved in Embden-Meyerhof pathway (EMP).**

| Gene | Ratio[a] | | | | Expression with IPTG addition[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | 18 h[c] | 24 h[c] | 30 h[c] | 36 h[c] | 18 h | 24 h | 30 h | 36 h |
| Gluconeogenic | | | | | | | | |
| *ppsA* | 7.4 | 10 | nc | nc | 81 | 34 | 17 | 13 |
| *pckA* | 0.7 | 0.7 | 0.4 | 0.4 | 0.9 | 0.5 | 0.5 | 0.5 |
| *adk* | 0.6 | 0.4 | 0.4 | 0.3 | 3.3 | 2.6 | 0.6 | 0.2 |
| Embden-Meyerhof pathway | | | | | | | | |
| *pgi* | nc | 1.9 | nc | 0.5 | 6.7 | 2.2 | 0.7 | 0.5 |
| *pfkA* | 1.2 | 1.3 | 1.2 | 0.7 | 2.7 | 1.8 | 1.8 | 1.1 |
| *pfkB* | 1.5 | 1.7 | 2.4 | 1.8 | 4.2 | 1.9 | 4.8 | 3.2 |
| *fbaB* | 1.1 | 1.9 | nc | 0.7 | 2.9 | 4.3 | 2.3 | 1.7 |
| *fbaA* | nc | 1.4 | 1.3 | nc | 6.8 | 3.9 | 2.3 | 2.1 |
| *gapA* | nc | 1.7 | 1.4 | 1.6 | 12 | 5.7 | 2.2 | 1.8 |
| *pgk* | nc | 2.2 | nc | nc | 14 | 9.9 | 2.4 | 2.2 |
| *tpiA* | 1.3 | 1.5 | nc | nc | 4.2 | 2.2 | 1.0 | 1.0 |
| *epd* | nc | nc | nc | 1.2 | 2.1 | 2.3 | 3.3 | 3.0 |
| *pgmI* | 1.3 | 1.3 | nc | nc | 3.0 | 3.2 | 2.2 | 1.7 |
| *gpmA* | 1.2 | 1.1 | nc | 0.8 | 15 | 12 | 11 | 11 |
| *gpmB* | nc | 1.3 | nc | nc | 1.6 | 1.6 | 1.0 | 0.8 |
| *eno* | 1.6 | 3.5 | 1.6 | 2.2 | 30 | 11 | 4.9 | 6 |
| *pykA* | nc | nc | nc | 0.8 | 2.1 | 1.5 | 1.8 | 1.2 |
| *pykF* | nc | nc | nc | nc | 6.0 | 3.1 | 3.0 | 1.5 |

[a] Expression ratio was calculated as (+IPTG/-IPTG). [b] percent cDNA $\times 10^2$.
[c] nc, no change based on student t-test ($p > 0.05$).

kinase, *tpiA* encoding triose phosphate isomerase, *eno* encoding enolase, *pykF* encoding pyruvate kinase, expression levels declined to 10-30% of initial values during the later stages of the fermentation both with IPTG addition and without IPTG addition (data not shown).

Most glycolytic genes were thought to be unregulated, with expression levels varying less than 2-fold.[14] However, many genes involved in glycolysis were judged to express at higher levels in *E. coli* KL3/pJY1.216A with PEP synthase overexpression during early fermentation (18 h and 24 h). Increases in expression levels of 12 genes in Embden-Meyerhof pathway were significant ($p < 0.05$) at 24 h as shown in Table 15.

Two genes (*eno* encoding enolase and *pfkB* encoding 6-phosphofructokinase) were significantly higher with PEP synthase overexpression at all time points.

However, even most of the glycolytic enzymes have been overexpressed at individual time points, the average glycolytic flux during the fermentation stays the same as measured by the total amount of glucose consumed during the fermentation. As suggested by Koebmann[15] and Ingram,[16] the control of glycolytic flux resides not inside but outside the pathway (demand for ATP), i.e., with the enzymes that hydrolyze ATP. Because PEP synthase uses ATP to synthesize PEP, the ATP demand should be higher for PEP synthase overexpression strain. It appeared that the ATP demand controls the expression level of glycolytic enzymes but has no control over the glycolytic flux in the system examined here. A possible explanation for this lack of stimulation of the glycolytic flux could be that glycolysis is already running close to its maximum capacity under the high density fed-batch fermentation conditions. A second possibility is that there are still rate-limiting enzymes in the pathway that have not been overexpressed, for instance in the transport of sugar into the cell or in the reactions that consume the products of glycolysis.

**Pentose pathway**

Table 16 summarizes the expression data for 10 genes involved in both oxidative and non-oxidative pentose pathway. Although results at each time point were normalized for total bound cDNA, expression levels of 7 genes stay almost the same at all the stages of the fermentation. Higher values were observed for three genes (*tktA* encoding transketolase, *gnd* encoding 6-phosphogluconate dehydrogenase and *pgl* encoding 6-

phosphogluconolactonase) during early fermentation (18 h and 24 h) than during the later stages of fermentation (30 h and 36 h). The *gnd* gene encoding 6-phosphogluconate dehydrogenase is known to be under the control of the growth rate.[17] The change of growth rate corresponds to the change of expression level of *gnd* gene. One particular interesting observation is that transketolase expression levels dropped down approximately 10 fold during the later stages of fermentation (30 h and 36 h). One possible explanation could be that the transketolase promoter is downregulated during the later stages of fermentation. However, the transketolase specific activities during the fermentation were pretty stable during the fermentation (approximately 0.02 U/mg, data not shown).

Most pentose pathway genes did not vary significantly with the overexpression of the PEP synthase. Two exceptions are *talA* and *talB* encoding two isozymes of transaldolase, which are known to enhance E4P supply.[18] Two transaldolase genes were significantly higher with PEP synthase overexpression. The increase in expression levels of transaldolase genes with overexpression of PEP synthase would be expected to result in higher specific activities for transaldolase and may provide a physiological basis for the higher rates of E4P supply relative to the rates of E4P supply with wild type expression of PEP synthase. As higher rates of E4P supply coupled with higher rates of PEP supply in "high yield" cells, more carbon flow will be expected to direct into the shikimate pathway. However, the possibility that E4P supply is still limiting the carbon flow to the shikimate pathway with amply supply of PEP cannot be ruled out.

Another interesting observation is that although the expression level of *pgi* encoding phosphoglucose isomerase, the first enzyme in the Embden-Meyerhof pathway,

declined to 10% of initial values during the later stages of the fermentation (Table 15), the expression level of *zwf* encoding glucose-6-phosphate dehydrogenase, the first

**Table 16. Expression data for genes involved in pentose pathway.**

| Gene | Ratio[a] | | | | Expression with IPTG addition[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | 18 h[c] | 24 h[c] | 30 h[c] | 36 h[c] | 18 h | 24 h | 30 h | 36 h |
| *ppsA* | 7.4 | 10 | nc | nc | 10 | 34 | 17 | 13 |
| Pentose pathway | | | | | | | | |
| *talA* | nc | 2.3 | 1.4 | nc | 5.8 | 4 | 4.3 | 3.4 |
| *talB* | 1.2 | 1.3 | 3.1 | 1.6 | 2.1 | 2.1 | 3.5 | 1.8 |
| *zwf* | nc | nc | nc | 1.7 | 2.0 | 1.5 | 2.1 | 1.6 |
| *gnd* | nc | nc | 1.4 | nc | 4.7 | 5.0 | 2.1 | 1.5 |
| *rpe* | 1.4 | nc | nc | nc | 1.0 | 0.6 | 1.9 | 2.2 |
| *rpiA* | 0.9 | 1.2 | nc | nc | 2.1 | 1.4 | 1.0 | 0.9 |
| *tktA* | nc | nc | 0.63 | 0.44 | 320 | 198 | 47 | 28 |
| *tktB* | 1.3 | nc | 1.5 | 2.6 | 7.9 | 3.4 | 7.3 | 7.6 |
| *rpiB* | nc | nc | nc | nc | 0.1 | 0.2 | 0.3 | 0.4 |
| *pgl* | nc | 1.9 | nc | 0.5 | 6.7 | 2.2 | 0.7 | 0.6 |

[a] Expression ratio was calculated as (+IPTG/-IPTG). [b] percent cDNA $\times 10^2$.
[c] nc, no change based on student t-test ($p > 0.05$).

enzyme in the oxidative pentose pathway, remained the same at all stages of the fermentation (Table 16). Since the ratio of Zwf and Pgi determine the carbon distribution between Embden-Meyerhof pathway and pentose pathway, the transcription result suggested that the oxidative pentose pathway played a more important role for the metabolism of glucose during the later stages of the fermentation than the Embden-Meyerhof pathway.

## TCA Cycle

Table 17 summarizes the expression data for all 24 genes involved in the TCA cycle. Although results at each time point were normalized for total bound cDNA, higher values were observed for most of these genes during early fermentation (18 h and 24 h) than during the later stages of fermentation (30 h and 36 h). For many genes, expression levels declined to 10-30% of initial values (18 h) during the later stages of the

fermentation. Only three genes (*aceK* encoding isocitrate dehydrogenase phosphatase, *fumB* encoding fumarase, *sucC* encoding succinyl-CoA synthetase) appeared to be stable during the fermentation.

**Table 17. Expression data for all genes involved in TCA cycle.**

| Gene | Ratio[a] | | | | Expression with IPTG addition[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | 18 h[c] | 24 h[c] | 30 h[c] | 36 h[c] | 18 h | 24 h | 30 h | 36 h |
| *ppsA* | 7.4 | 10 | nc | nc | 10 | 34 | 17 | 13 |
| TCA cycle | | | | | | | | |
| *gltA* | 0.8 | 1.3 | 1.4 | 1.5 | 21 | 9.6 | 4.5 | 3.1 |
| *acnB* | nc | 1.7 | 2.2 | nc | 3.4 | 2 | 1.3 | 0.7 |
| *acnA* | nc | 5.9 | nc | nc | 8.7 | 6.7 | 2 | 2.5 |
| *icdA* | 0.7 | 2 | 2.2 | 1.6 | 14 | 13 | 4.7 | 3.2 |
| *aceA* | 0.9 | 4.7 | nc | nc | 17 | 10 | 1.4 | 1.8 |
| *aceB* | nc | 4 | 2.6 | 2.3 | 23 | 11 | 3.5 | 2.9 |
| *aceK* | 1.3 | nc | nc | nc | 3.5 | 1.8 | 2.3 | 4.1 |
| *mdh* | 1.5 | nc | 1.8 | nc | 15 | 8 | 6.4 | 2.9 |
| *ppc* | nc | 1.7 | 1.8 | nc | 10 | 2.4 | 1.6 | 1 |
| *fumA* | nc | 1.4 | 1.3 | 0.4 | 1.6 | 1.1 | 1 | 0.4 |
| *fumB* | 1.7 | nc | nc | 2.3 | 0.8 | 0.3 | 1.3 | 1.9 |
| *fumC* | nc | 5.7 | nc | 0.6 | 1.2 | 3.3 | 1 | 0.9 |
| *sdhD* | 0.8 | 2.1 | nc | nc | 1.7 | 0.5 | 0.2 | 0.4 |
| *sdhC* | 0.7 | 1.9 | nc | nc | 3.1 | 1.2 | 0.7 | 0.4 |
| *sdhB* | 0.8 | 2.1 | nc | nc | 5.4 | 2 | 0.9 | 0.6 |
| *sdhA* | 0.8 | 1.3 | nc | nc | 6.6 | 3.5 | 1.3 | 1 |
| *sucC* | 0.9 | 1.8 | nc | 7.8 | 9.3 | 2.3 | 6.3 | 9.7 |
| *sucD* | nc | 3.6 | nc | nc | 9.9 | 3 | 0.9 | 0.7 |
| *sucB* | 0.8 | 2.3 | nc | 0.6 | 6.1 | 1.6 | 0.6 | 0.5 |
| *sucA* | nc | 3.7 | nc | 0.3 | 11 | 2.7 | 0.4 | 0.4 |
| *mqo* | nc | nc | 0.3 | 0.1 | 1.1 | 1.6 | 0.5 | 0.3 |
| *glcB* | nc | nc | nc | 0.1 | 0.5 | 0.1 | 0.1 | 0.1 |
| *prpC* | nc | nc | nc | 0.4 | 1.5 | 1.4 | 0.6 | 0.6 |

[a] Expression ratio was calculated as (+IPTG/-IPTG). [b] percent cDNA $\times 10^2$.
[c] nc, no change based on student t-test ($p > 0.05$).

At 18 h (exponential growth phase), some genes involved in the TCA cycle were downregulated (0.7-0.9) while at 24 h (early stationary phase) most of the genes involved in TCA cycle were upregulated by 2- to 6- fold in *E. coli* KL3/pJY1.216A with PEP synthase overexpression. The glyoxylate shunt genes (*aceB* encoding malate synthase, *aceA* encoding isocitrate lyase) were significantly higher with PEP synthase overexpression at most of the time points. Because carbon loss to carbon dioxide can be

avoided using glyoxylate shunt, more usage of this pathway can lead to higher product yield.

**Table 18. Expression data for global regulators.**

| Gene | Ratio[a] | | | | Expression with IPTG addition[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | 18 h[c] | 24 h[c] | 30 h[c] | 36 h[c] | 18 h | 24 h | 30 h | 36 h |
| *ppsA* | 7.4 | 10 | nc | nc | 10 | 34 | 17 | 13 |
| Regulators | | | | | | | | |
| *arcA* | 0.9 | nc | 0.7 | 0.6 | 4.9 | 6.5 | 2.7 | 1.6 |
| *arcB* | nc | 1.3 | nc | 0.3 | 3.6 | 1.6 | 0.7 | 0.4 |
| *crp* | 1.2 | nc | nc | nc | 0.7 | 1.0 | 0.9 | 0.7 |
| *cyaA* | nc | 2.1 | 4.0 | 1.8 | 1.8 | 1.0 | 2.8 | 1.6 |
| *fadR* | 0.7 | 0.8 | 0.7 | 0.7 | 5.5 | 5.3 | 3.1 | 2.0 |
| *iclR* | nc | 0.9 | nc | 1.5 | 1.8 | 2.4 | 2.3 | 4.1 |
| *fruR* | nc | nc | nc | 1.4 | 1.0 | 0.9 | 1.9 | 2.1 |
| *csrA* | nc | 1.3 | 0.5 | 0.5 | 2.2 | 1.9 | 1.8 | 0.9 |
| *fnr* | nc | nc | 0.8 | nc | 2.6 | 3.1 | 2.4 | 2.4 |

[a] Expression ratio was calculated as (+IPTG/-IPTG). [b] percent cDNA $\times 10^2$.
[c] nc, no change based on student t-test ($p > 0.05$).

The up-regulation of most TCA cycle genes during the stationary phase was expected since overexpression of PEP synthase consumed a lot of ATP that must be compensated by the TCA cycle. These TCA cycle genes are known to be regulated by several regulators, such as cAMP-CRP, ArcA, and Fnr.[19] From the expression data for those control genes (Table 18), it is likely that the observed up-regulation of TCA cycle genes with PEP synthase overexpression was mediated by cAMP-CRP, as cAMP level might be increased with PEP synthase overexpression because the *cyaA* gene encoding adenylate cyclase that is responsible for the synthesis of cAMP was upregulated 2- to 4-fold with PEP synthase overexpression during the stationary phase (24 h, 30 h and 36 h). For the glyoxylate shunt genes (*aceA* and *aceB*), the downregulation of FadR repressor protein may lead to their upregulation.

**Table 19. Expression data for genes involved in the shikimate pathway.**

| Gene | Ratio[a] | | | | Expression with IPTG addition[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | 18 h[c] | 24 h[c] | 30 h[c] | 36 h[c] | 18 h | 24 h | 30 h | 36 h |
| *ppsA* | 7.4 | 10 | nc | nc | 10 | 34 | 17 | 13 |
| Shikimate pathway | | | | | | | | |
| *trpA* | 0.9 | nc | 5.7 | 2.9 | 1.6 | 1 | 12 | 6.2 |
| *trpB* | nc | 2.8 | 10 | 4 | 0.2 | 0.3 | 1.5 | 0.6 |
| *trpC* | nc | nc | 7.6 | 3.8 | 0.9 | 0.6 | 22 | 12 |
| *trpD* | nc | 1.5 | 7.2 | 3.2 | 1.1 | 0.8 | 32 | 14 |
| *trpE* | 1.4 | nc | 3.8 | 2.5 | 0.7 | 0.3 | 43 | 24 |
| *aroF* | nc | 0.5 | 0.7 | nc | 262 | 549 | 780 | 1070 |
| *aroG* | 0.9 | 0.4 | 0.5 | 1.8 | 10 | 6.0 | 23 | 21 |
| *aroH* | nc | nc | nc | 0.5 | 0.4 | 0.2 | 0.8 | 0.4 |
| *aroB* | nc | 0.9 | nc | 0.8 | 5.6 | 6.5 | 6.3 | 5.7 |
| *aroD* | nc | nc | nc | 0.3 | 1.0 | 0.5 | 0.3 | 0.2 |
| *aroK* | nc | 1.3 | nc | 0.4 | 2.6 | 1.8 | 0.9 | 0.5 |
| *aroL* | 0.7 | nc | nc | 0.7 | 0.9 | 1.6 | 1.6 | 1.1 |
| *aroA* | nc | 1.3 | nc | nc | 3.2 | 2.4 | 2.4 | 2.8 |
| *aroC* | 0.9 | 1.7 | nc | 0.4 | 1.7 | 1.0 | 0.4 | 0.4 |
| *aroE* | nc | nc | nc | 0.5 | 2.2 | 2.1 | 2.1 | 1.2 |
| *pheA* | 0.5 | nc | 1.4 | 0.6 | 2.0 | 1.5 | 1.1 | 0.7 |
| *tyrA* | nc | 0.8 | nc | nc | 9.7 | 27 | 35 | 32 |

[a] Expression ratio was calculated as (+IPTG/-IPTG). [b] percent cDNA $\times 10^2$.
[c] nc, no change based on student t-test ($p > 0.05$).

## Shikimate pathway

Table 19 summarizes the expression data for 17 genes involved in the shikimate pathway. The total amount of *aroF* cDNA was a large portion of the total cDNA. At 36 h, it represents about 10% of the total cDNA, indicating that the promoter $P_{aroF}$ is very strong during the later stages of the fermentation. With PEP synthase overexpression, the continuous increase of expression level of $aroF^{FBR}$ during the course of the fermentations did not lead to the continuous increase of DAHP synthase specific acitivites during the course of the fermentations (Figure 31). This result suggests that $AroF^{FBR}$ may be sensitive to proteolysis during the stationary phase. The expression level of *aroB* encoding 3-dehydroquinate synthase was very stable during the course of the fermentations both with IPTG addition and without IPTG addition. Thus the promoter of *aroB* may be used as a weak promoter to express certain genes in the future.

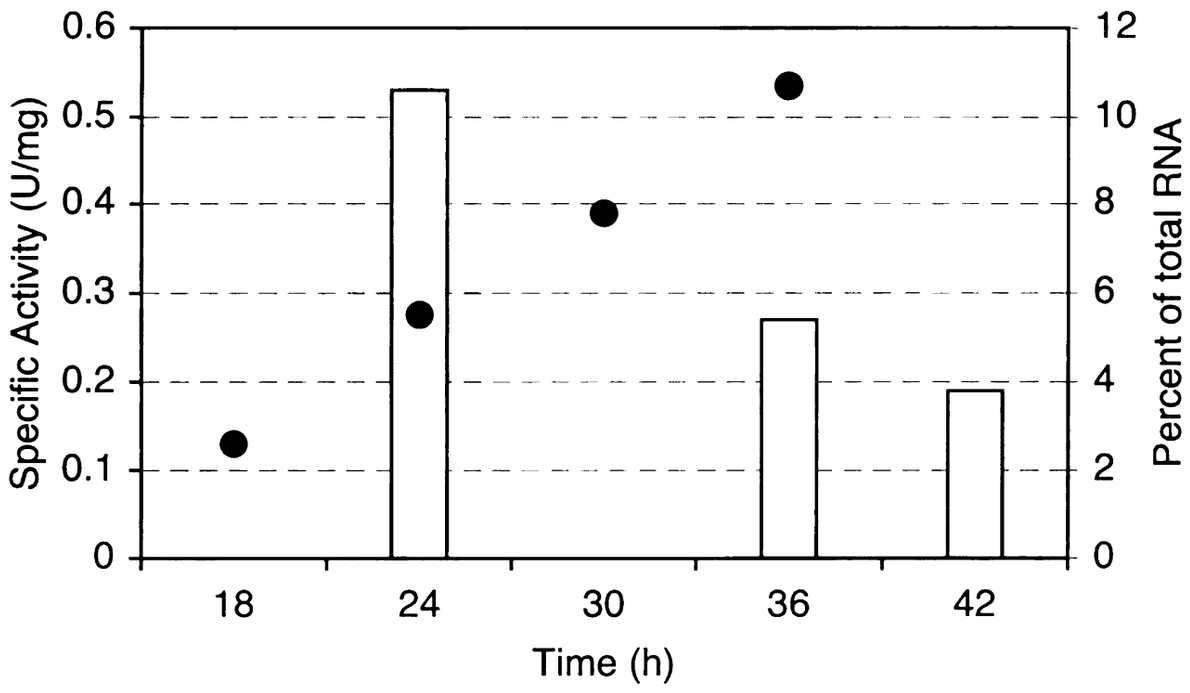



**Figure 31. The specific activity of AroF$^{FBR}$ (open bars) and *aroF*$^{FBR}$-encoded mRI transcript (filled circles) over the course of an *E. coli* KL3/pJY1.216A fermentor r with IPTG addition.**

For the genes involved in the biosynthesis of tryptophan, tyrosine a phenyalanine (*trpABCDE*, *tyrA*, *pheA)*, expression levels increase to approximately 4- 20- fold of initial values (18 h) during the later stages of the fermentation (30 h and 36 indicating the cell's need for those three aromatic amino acids after the initial addition those three aromatic amino acids has been consumed. Phenylalanine and tyrosine w consumed completely before tryptophan was consumed at 24 h. The *tyrA* gene encod chorismate mutase that leads to the biosynthesis of tyrosine and phenylalanine has be induced 3 fold while the *trpA* gene encoding tryptophan synthase that leads to biosynthesis of tryptophan had not been induced at 24 h.

It is surprising to see that in high yield cells, *aroF*$^{FBR}$ encoding the first enzyme the shikimate pathway was downregulated at 24 h and 30 h compared to low yield ce However, this downregulation can be an effective way to reduce the metabolic burden the cells with overexpressing both DAHP synthase and PEP synthase. Genes in

tryptophan pathway (*trpABCDE*) were unexpectedly upregulated 1.4- to 10- fold in high yield cells during the later stages of the fermentation (30 h and 36 h). The underlying mechanism for this upregulation was unknown.

**Table 20. Expression data for selected genes related to stress conditions.**

| Gene | Ratio[a] | | | | Expression with IPTG addition[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | 18 h[c] | 24 h[c] | 30 h[c] | 36 h[c] | 18 h | 24 h | 30 h | 36 h |
| *ppsA* | 7.4 | 10 | nc | nc | 10 | 34 | 17 | 13 |
| Molecular chaperone | | | | | | | | |
| *dnaK* | nc | 2.6 | 3.2 | 3.1 | 13 | 5.9 | 5.6 | 5.9 |
| *dnaJ* | nc | nc | 1.8 | 1.4 | 3.3 | 1.8 | 5.4 | 4.5 |
| *mopB* | 0.9 | 7.5 | 2.6 | 1.6 | 4.7 | 2.2 | 0.8 | 0.6 |
| *grpE* | nc | 1.2 | 1.3 | nc | 2.8 | 2.2 | 2.0 | 2.0 |
| *clpB* | nc | 1.5 | 1.9 | 2 | 4.4 | 2.6 | 2.6 | 3.5 |
| Protease | | | | | | | | |
| *lon* | nc | nc | nc | nc | 1.5 | 1.2 | 1.2 | 1.0 |
| Sigma Factor | | | | | | | | |
| *rpoE* | 1.1 | 2.5 | nc | 2.1 | 1.3 | 0.9 | 0.9 | 2.5 |
| *rpoH* | 0.7 | nc | 1.8 | 1.3 | 2.9 | 3.8 | 11 | 6.5 |
| Oxidative stress | | | | | | | | |
| *katE* | 1.8 | 3.1 | nc | nc | 8.9 | 4.0 | 1.7 | 1.7 |
| *katG* | nc | nc | nc | nc | 2.5 | 1.8 | 1.0 | 1.7 |
| *sodA* | nc | 4.1 | 2.3 | 2.3 | 4.8 | 2.5 | 1.6 | 2.2 |
| *sodB* | nc | nc | 0.8 | 1.3 | 6.1 | 4.1 | 4.1 | 4.1 |
| *gldA* | 1.4 | 2.1 | 3.5 | 2.5 | 1.2 | 1.9 | 4.2 | 2.4 |
| *poxB* | 1.5 | 7.0 | 3.6 | 2.4 | 4.6 | 2.3 | 1.8 | 1.9 |
| *b1005* | 1.6 | 5.7 | 3.3 | 2.2 | 1.8 | 2.8 | 1.9 | 2.3 |
| Osmotic stress | | | | | | | | |
| *kdpA* | nc | nc | 4.3 | 18 | 1.0 | 0.4 | 28 | 32 |
| Acid stress | | | | | | | | |
| *gadA* | 1.8 | 1.7 | nc | 0.5 | 1.3 | 0.7 | 0.4 | 0.4 |
| *gadB* | 2.3 | nc | 0.6 | 0.6 | 4.5 | 1.3 | 1.3 | 1.2 |
| *xasA* | 3.4 | nc | nc | 0.3 | 0.9 | 0.2 | 0.2 | 0.2 |
| *hdeA* | 4.7 | 3.3 | 0.5 | 0.3 | 8.3 | 1.6 | 0.4 | 0.2 |
| *hdeB* | 4.5 | 4.4 | 0.4 | 0.1 | 4.7 | 0.6 | 0.1 | 0.1 |
| *marA* | nc | 2.8 | 1.3 | 1.3 | 7.2 | 15.5 | 3.3 | 3.3 |
| *marB* | nc | 2.0 | nc | nc | 1.4 | 2.0 | 1.0 | 1.0 |
| DNA protection | | | | | | | | |
| *dps* | 1.4 | 5.3 | 2 | 1.8 | 7.0 | 19 | 24 | 29 |

[a] Expression ratio was calculated as (+IPTG/-IPTG). [b] percent cDNA $\times 10^2$.
[c] nc, no change based on student t-test ($p > 0.05$).

## Genes related to stress conditions

A general function of the heat shock protein is to monitor and respond to the state of protein folding in the cell. Many of the heat shock proteins are molecular chaperones, whose function is to bind to newly synthesized, partially folded or unfolded proteins and

promote their folding and refolding by limiting the nonproductive interactions that lead aggregation or misfolding. Other heat shock proteins are proteases that function degrade misfolded or abnormal proteins. From this point of view, it is qui understandable that the heat shock proteins are required for cell viability when the cel are under stress in the high cell density culture.

Table 20 summarizes the expression data for selected genes related to stre conditions. In general, major heat shock genes (*dnaKJ, grpE, mopB*) were upregulate in high yield cells during the stationary phase (24 h, 30 h and 36 h) than the low yie cells. It may reflect the fact that more proteins in high yield cells were produced than low yield cells. However, the expression level of major protease gene *lon* was unchange in both high yield and low yield cells. The increased expression levels of *rpoE and rop* (encoding $\sigma^E$ and $\sigma^{32}$, respectively) may be responsible for the increased expression leve of major heat shock genes.

The *sodA* gene encoding superoxide dismutase and the *katE* gene encodin hydroperoxidase II that are related to peroxide damage were expressed at significant higher value (2-4 fold) in high yield cells at individual time point. This may indicate th existence of higher concentration of superoxide inside the high yield cells. In literatu the *sodAsodB* strain of *E. coli* exhibited an auxotrophy for three aromatic amino acid (phenylalanine, tyrosine and tryptophan) and superoxide was show[20] to interfere with th production of erythrose-4-phosphate by oxidizing an intermediate in the transketola reaction with a rate constant of ~$10^6$ $M^{-1}S^{-1}$.[21] The upregulation of the superoxid removing genes (*sodA and katE*) in high yield cells indicated a very important role superoxide protection for the biosynthesis of aromatic amino acids. Pyruva

dehydrogenase (*poxB*), glycerol dehydrogenase (*gldA)* and a hypothetic NADPH dependent dehydrogenase (*b1005*) were also expressed at significantly higher value in high yield cells. These genes are both effective to resist the oxidative stress in *E. coli* cells.

A very important gene related to osmotic stress, *kdpA* encoding potassium transporting ATPase was expressed at higher levels (4-fold to 18-fold) in high yield cells during the later stages of the fermentation (30 h, 36 h). The most rapid response to osmotic upshock is a stimulation of uptake of $K^+$. KdpA is an inducible system that serves to scavenge $K^+$ when the ion is present at low concentrations.[22] The higher expression level of *kdpA* gene may indicate the higher osmolarity inside the high yield cells than the low yield cells.

Glutamate decarboxylase (*gadA* and *gadB*) and the glutamate transporter (*xas*) are particularly important for acid tolerance and the maintenance of internal pH. Expression of these genes is related to both pH and the abundance of organic acids.[23] Expression levels for *gadA* and the *gadB-xasA* operon were 2-4 fold higher with PEP synthase overexpression during early fermentation at 18 h and 24 h, consistent with the production of higher levels of organic acids with PEP synthase overexpression. Some other genes related to acid stress such as acid-resistant unknown genes (*hdeA*, *hdeB*), multidrug resistance genes (*marA, marB*) were also expressed higher in high yield cells with PEP synthase overexpression during early fermentation at 18 h and 24 h.

Almiron et al.[24] identified a DNA-binding protein, designated Dps (DNA-binding protein from starved cells), which is produced primarily in stationary-phase cells of *E. coli*. Dps forms spherical dodecamers, homologous to ferritins, which sequester and

protect DNA from oxidative stress, nuclease, and UV light. Dukan and Touati[25] showe that mutations of *recA*, *recB*, and *dps* in *E. coli* rendered cells more sensitive to damag from hydroxyl radicals generated by HOCl. This suggests that not only DNA repair, bu also DNA protection by *dps*, is pivotal for survival in extreme conditions. In high yiel cells with PEP synthase overexpression, *dps* gene was expressed at significant highe levels during the course of fermentation relative to the low yield cells. This resul confirmed with the above expression data that the high yield cells were under highe stress (acid stress and oxidative stress) than the low yield cells.

**Transport Genes**

Table 21 summarizes the expression data for some selected transport genes. Th glucose transport genes *ptsHI*, *crr* were all upregulated in high yield cells with PE synthase overexpression compared to in low yield cells. However, the total amount o glucose consumed during the fermentation did not change. At later stages of th fermentation (30 h and 36 h), many PTS transport genes (*ptsN* encoding unknown PT enzyme II that regulates N metabolism, *cmtB* encoding PTS enzyme II for mannitol, *srl* encoding PTS enzyme II for sorbitol, *manZ* encoding PTS enzyme II for mannose) an other transport genes (*b2968* encoding unknown transport protein, *bcp* encodin bacterioferritin comigratory protein, *bfr* encoding bacterioferritin storage protein, *potI* encoding putrescine ABC transporter, *kdpA* encoding potassium-transporting ATPase were up-regulated 2- to 70- fold in high yield cells compared to in low yield cells. Th upregulation of PTS transport genes may reflect the cell's willingness to find mor nutrients when PEP was abundant inside the cells.

**Table 21. Expression data for selected transport genes.**

| Gene | Ratio[a] | | | | Expression with IPTG addition[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | 18 h[c] | 24 h[c] | 30 h[c] | 36 h[c] | 18 h | 24 h | 30 h | 36 h |
| *ppsA* | 7.4 | 10 | nc | nc | 10 | 34 | 17 | 13 |
| PTS transport | | | | | | | | |
| *ptsH* | nc | 1.8 | 2.3 | 4.6 | 4.1 | 1 | 2.7 | 2.3 |
| *ptsI* | nc | 1.7 | 2.7 | 1.4 | 10 | 3.4 | 4.9 | 2.3 |
| *srlA* | nc | nc | nc | 10 | 0.4 | 0.1 | 2.2 | 6 |
| *crr* | nc | 2.2 | 2.6 | nc | 2.6 | 1.8 | 2.0 | 0.7 |
| *ptsN* | 1.4 | nc | 4.1 | 73 | 4 | 2 | 63 | 79 |
| *cmtB* | nc | 0.4 | 1.9 | 15 | 0.3 | 0.5 | 4 | 7.1 |
| *manZ* | nc | nc | 1.7 | 4.4 | 2.2 | 0.6 | 3.2 | 5.7 |
| *ptsG* | nc | nc | nc | nc | 0.3 | 0.1 | 0.2 | 0.1 |
| *glvC* | 0.9 | nc | nc | 3.4 | 3.0 | 3.0 | 4.3 | 7.5 |
| Other transport | | | | | | | | |
| *kdpA* | nc | nc | 4.3 | 18 | 1.0 | 0.4 | 28 | 32 |
| *b2968* | nc | nc | 2.1 | 11 | 1.9 | 1.2 | 13 | 30 |
| *bcp* | nc | nc | 2.6 | 17 | 1.4 | 1.3 | 6.5 | 11 |
| *bfr* | nc | 2.6 | 3.2 | 2.6 | 1.8 | 1.6 | 3.4 | 1.6 |
| *potH* | 1.4 | 0.6 | 2.6 | 5.6 | 1.0 | 0.9 | 5.2 | 5.9 |

[a] Expression ratio was calculated as (+IPTG/-IPTG). [b] percent cDNA $\times 10^2$.
[c] nc, no change based on student t-test ($p > 0.05$).

## Conclusion

This work demonstrates the utility of the DNA microarray technology in studying gene expression profiles under fed-batch fermentor conditions. The results showed that a simple overexpression of PEP synthase led to dramatic expression changes of many genes. Some interesting expression changes happened in the TCA cycle, pentose pathway, Embden-Meyerhof pathway, shikimate pathway, cell processes (resistance to oxidative stress, acidic stress and osmotic stress) and transport proteins.

This comparison between high yield cells and low yields cells shows that transcript levels may not be used to infer fluxes in a quantitative manner. For example, the increase of expression levels of many glycolysis genes at 24 h and 30 h does not correspond to the increase of glycolytic flux as measured by the amount of glucose consumed during the course of the fermentation (data not shown). However, changes in transcript levels reveal the existence of significant regulation and suggest possible

strategies of genetic modification. This expression profiles, combined with previous knowledge of metabolic pathway and fermentation results, identified many interesting gene targets for further strain development. For example, the large amount of *aroF*$^{FBR}$ cDNA suggests the degradation of AroF$^{FBR}$ during the stationary phase of the fermentation. The downregulation of *tktA* gene during the later stages of the fermentation suggests that the availability of E4P may be the limiting factor for driving carbon flow to shikimate pathway. The upregulation of *sodA* gene in high yield cells suggests that superoxide accumulated during the fermentation may be a limiting factor for driving the carbon flow to shikimate pathway because superoxide can blocking an intermediate in the transketolase reaction.[20] The upregulation of some unknown transport genes such as *ptsN* and *b2968* at 36 h in high yield cells generates interests about their functions. In the following section, two proposed experiments toward enhancing biocatalyst development will be discussed and the first proposed experiment was tested.

## Experiments toward enhancing biocatalyst development

### Feedback insensitive isozyme of DAHP synthase

The activity of DAHP synthase has attracted attention as a critical determinant of carbon flow directed into the shikimate pathway.[26] Given the importance of feedback inhibition in determining the activity of DAHP synthase in the microbial cytoplasm, use of mutant isozymes of DAHP synthase that are insensitive to feedback inhibition by aromatic amino acids has been a centerpiece of virtually all efforts to synthesize shikimate pathway intermediates and products in high yield. A density gradient chamber

was used to isolate an isozyme of tyrosine-sensitive DAHP synthase that had reduced sensitivity to feedback inhibition by aromatic amino acids.[27] The gene encoding this feedback-insensitive DAHP synthase, which was designated *aroF*$^{FBR}$, corresponded to a feedback-insensitive isozyme of DAHP synthase previously reported by Herrmann.[26b]

Although *aroF*$^{FBR}$-encoded DAHP synthase has been our workhorse in studies of hydroaromatics synthesis, we routinely observed that the specific activity of DAHP synthase declined over the course of fermentor runs (Figure 31).[28, 1a, 1c] What became even more alarming were the results from microarray analysis described above. These experiments revealed (Figure 31) a decline in DAHP synthase activity but an increase in the *aroF*$^{FBR}$ transcript over the course of fermentor runs. By the end of the fermentation, *aroF*$^{FBR}$ transcript accounted for over 10% of the total transcriptome. This observation raises the possibility that AroF$^{FBR}$ declined in specific activity even though transcription and translation of this gene may have taken place over the entire course of the fermentation runs. AroF$^{FBR}$ might be even more sensitive to proteolysis than indicated by the declined specific activities of this enzyme over the course of a fermentation run.

The literature suggests that phenylalanine-sensitive AroG is more resistant towards proteolysis than tyrosine-sensitive AroF although the degree of increased stability is open to question.[29] Herrmann has reported that AroF activity declines with a half-life of 3 $h^{-1}$ while under the same conditions, the specific activity of AroG remains unchanged.[29a] Pittard has reported a 5.5 $h^{-1}$ half-life for the specific activity of AroF and a 7 $h^{-1}$ half-life for the specific activity of AroG.[29b] Given our observations relative to the specific activity and transcription of *aroF*$^{FBR}$ and the literature concerning the proteolytic

lability of AroF relative to AroG, feedback insensitive isozyme of *aroG*-encoded DAHP synthase was isolated to compare with $AroF^{FBR}$.

**Table 22. Comparison of the amino acid changes and sensitivity to feedback inhibition of AroG and AroF Mutants.**

| Mutant | Amino acid change | Feedback inhibition[a] |
|---|---|---|
| AroG1 | Ser-211 → Tyr | 33% |
| AroG2 | Met-147 → Lys | 42% |
| AroG3 | Pro-150 → Thr | 53% |
| AroG4 | Val-221 → Met | 54% |
| $AroF^{FBR}$ | Pro-148 → Leu | 100% |

[a] Feedback inhibition = 100%- (the specific activity of DAHP synthase isozyme in a solution containing 4 mM Phe, 4 mM Tyr, and 4 mM Trp / the specific activity of DAHP synthase isozyme in the absence of any added amino acids).

PCR mutagenesis[30] was used to generate variants of aroG. The top proiority of mutagenic PCR is to introduce the various types of mutations in an unbiased fashion rather than to achieve a high overall level of amplification. The protocol for mutagenic PCR is derived from standard PCR conditions. The following changes were made to enhance the mutation rate: the $MgCl_2$ concentration was increased from 1.5 mM to 7 mM to stabilize noncomplementary pairs; 0.5 mM $MnCl_2$ was added to diminish the template specificity of the polymerase; the concentration of dCTP and TTP was increased to 1 mM to promote misincorporation; the amount of *Taq* polymerase was increased to 5 units to promote chain extension beyond positions of base mismatches. The *E. coli aroG* ORF from plasmid pJY6.119A (wild-type *aroG* ORF amplified from *E. coli* RB791 genomic DNA) was subject to PCR mutagenesis and the resulting DNA was ligated to pJF118EH. The resulting ligation mixture was introduced by transformation into *E. coli* CB734, which is deficient in DAHP synthase activity due to a Δ*aroF*Δ*aroG*Δ*aroH* phenotype.[31]

Growth of *E. coli* CB734 in the presence of 20 mM phenylalanine supplementation, which is adequate to prevent the growth of *E. coli* CB734/pJY6.119A, indicates an *aroG*$^{FBR}$ isozyme plasmid insert.[32] From the selection plates 13 colonies were picked up and then assayed for the DAHP specific activity in the presence of 4 mM phenylalanine, tyrosine and tryptophan. The concentrations of the three aromatic amino acids in the enzyme assay were chosen to mimic the concentrations of the three aromatic amino acids in the fermentation medium. Four feedback-insensitive isozymes (Table 22) have the highest activity in the presence of all three aromatic amino acids were sequenced. A Met-147-Ile and a Pro-150-Leu mutation have previously been reported to give rise to *aroG*$^{FBR}$ isozymes.[33] The AroG1 mutant (Ser-211-Tyr mutation), which has the highest resistance towards feedback inhibition, does not correspond to a position in AroG previously reported for an isozyme insensitive to feedback inhibition by phenylalanine. The AroG1 mutant also has almost 100% resistance towards feedback inhibition of each 4 mM aromatic amino acid (data not shown). One of the surprises in obtaining the *aroG*$^{FBR}$ isozymes is that our current *aroF*$^{FBR}$ displayed poor resistance to feedback inhibition at the higher concentration of aromatic amino acids used in these determinations (4 mM Phe, Tyr and Trp). In literature, the maximum concentration of the aromatic amino acids used in the examination of *aroF*$^{FBR}$ feedback inhibition was only 0.25 mM.[34]

**Biocatalyst design and fermentation**

To examine whether AroG1 is more stable over the course of fermentation runs than AroF$^{FBR}$, plasmid pJY6.231 (Figure 32) that carried *aroG*$^{FBR}$, $P_{aroF}$, *tktA* and *ppsA* inserts was assembled to mimic plasmid pJY1.216A (the best DHS producer). The *aroF*

promoter was used for expression of $aroG^{FBR}$ in order to minimize changes in

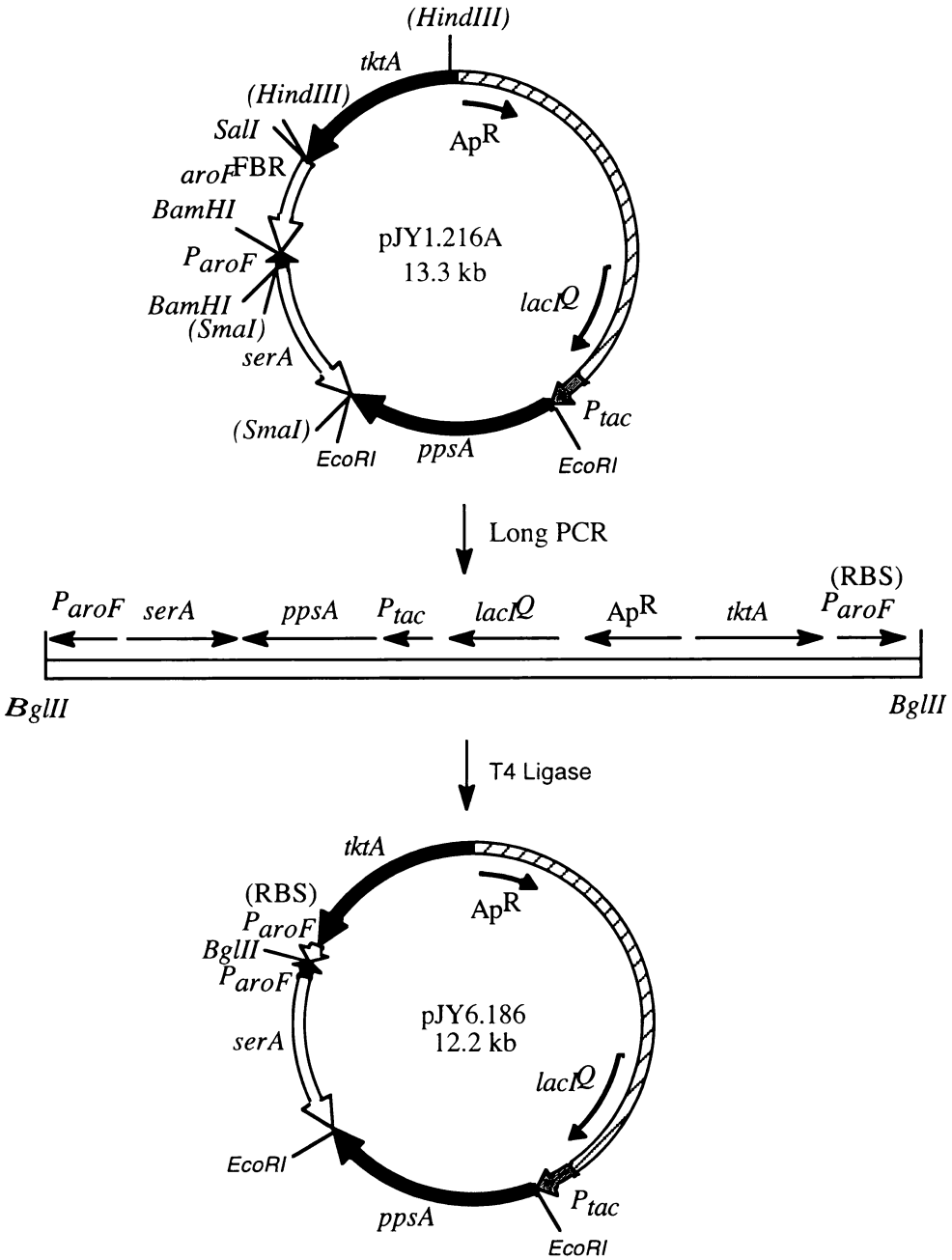



**Figure 32A. Construction of plasmid pJY6.186**

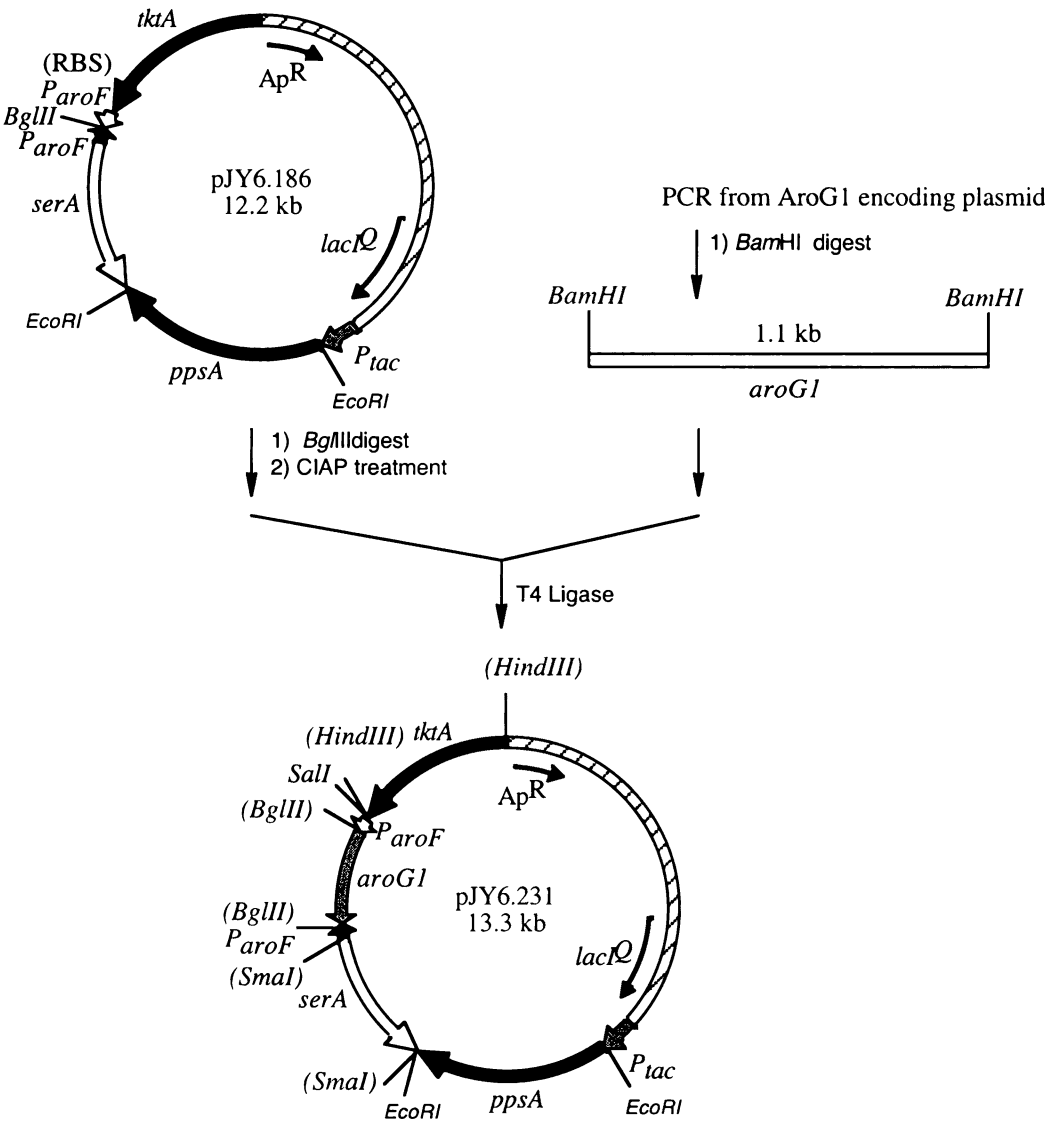



**Figure 32B. Construction of plasmid pJY6.231**

transcription levels. Both plasmids were transformed into *E. coli* strain KL3 and evaluated under fermentor-controlled glucose-rich conditions as described in Chapter 2. The specific activities of DAHP synthase during the fermentation of *E. coli* KL3/pJY6.231 (Table 23B) were not only stable but also higher than the DAHP synthase activities during the fermentation of *E. coli* KL3/pJY1.216A. However, no increase in titer and yield of DHS biosynthesis was observed (Table 23A). A reasonable explanation

for the observed ceiling in synthesized DHS concentrations with increasing DAHP synthase activity is that the availability of either intracellular PEP or intracellular E4P is still limiting even with the overexpression of both transketolase and PEP synthase.

**Table 23A. Concentrations and yields of 3-dehydroshikimic acid and byproducts synthesized by *E. coli* KL3/pJY1.216A and KL3/pJY6.231 cultured under glucose-rich conditions.**

| Construct | IPTG addn.[a] | [DHS][b] (g/L) | DHS Yield[c] (mol/mol) | [DAH] (g/L) | [DHQ] (g/L) | [GA] (g/L) | Total Yield[d] (mol/mol) |
|---|---|---|---|---|---|---|---|
| KL3/pJY6.231 | 12 | 64 | 35% | 13 | 9.2 | 9.5 | 50% |
| KL3/pJY1.216A | 12 | 69 | 35% | 10 | 12 | 13 | 51% |

[a]Amount of IPTG (mg) added at 18, 24, 30, and 36 h. [b]Abbreviations: 3-dehydroshikimic acid (DHS), 3-deoxy-D-*arabino*-heptulosonic acid (DAH), 3-dehydroquinic acid (DHQ), gallic acid (GA). [c](mol DHS)/(mol glucose consumed). [d](mol DHS + mol DAH + mol DHQ + mol GA)/ (mol glucose consumed).

**Table 23B. DAHP synthase specific activities for *E. coli* KL3/pJY1.216A and KL3/pJY6.231 cultured under glucose-rich conditions.**

| construct | IPTG addn.[b] | DAHP synthase (U/mg)[b] | | |
|---|---|---|---|---|
| | | 24 h | 36 h | 42 h |
| KL3/pJY6.231 | 12 | .87 | 0.96 | 1.0 |
| KL3/pJY1.216A | 12 | .53 | .27 | .19 |

[a]Amount of IPTG (mg) added at 18, 24, 30, and 36 h. [b]One unit (U) of DAHP synthase corresponds to formation of 1 μmole of DAHP per min at 37 °C.

In conclusion, a stable AroG$^{FBR}$ was obtained by PCR mutagenesis and has its advantage over AroF$^{FBR}$ because of its stability over the course of a fermentation run. Further study will include testing the stability of the *aroG*$^{FBR}$-encoded transcript and the *aroF*$^{FBR}$-encoded transcript over the course of a fermentation run using real-time reverse transcriptase-PCR (RT-PCR).[35] Real-time reverse transcriptase-PCR has become the method of choice for the rapid and quantitative examination of the expression of specific genes. Even small changes in the *aroG*$^{FBR}$-encoded transcript or *aroF*$^{FBR}$-encoded

transcript can be rapidly, reproducibly, and statistically determined using this method. Another advantage of real-time RT-PCR is that the integrity of the *aroG*$^{FBR}$-encoded transcript or *aroF*$^{FBR}$-encoded transcript can be estimated by using different pairs of primers in the real-time RT-PCR to quantitate RNA of different lengths. Also future study will include an assay to test the stability of the AroG$^{FBR}$ protein and AroF$^{FBR}$ protein over the course of the fermentation run. Before *aroG*$^{FBR}$ was incorporated into our constructs, it would be critical to determine whether there is an *aroG*$^{FBR}$ isozyme superior to our currently employed *aroF*$^{FBR}$.

**E-4-P limitation and superoxide protection**

The key determinant of carbon flow directed into a biosynthetic pathway is often the in vivo activity of the first enzyme in the pathway. For the common pathway of aromatic amino acid biosynthesis, the in vivo activity of DAHP synthase is dictated by feedback inhibition, transcriptional repression, and the availability of substrates E4P and PEP. In the system examined above (Tables 23A and 23B), the observed ceiling in synthesized shikimate pathway product concentrations with increasing DAHP synthase activity indicates that either intracellular PEP availability or intracellular E4P availability is still limiting the carbon flow to the shikimate pathway even with the overexpression of both transketolase and PEP synthase. The availability of PEP is not likely to be the limiting factor based on two observations. First, higher PEP synthase activity didn't correspond to high yields of shikimate pathway product as examined in Chapter 2.[1a] Second, the upregulation of many PTS transport proteins in high yield cells at late stationary phase (36 h) from the microarray analysis suggested there are still ample

supply of PEP at late stationary phase with the overexpression of PEP synthase. Thus the availability of E4P appeared to be the limiting factor under the fed-batch fermentor conditions employed. Then the question becomes why the supply of E4P is still limiting even with overexpression of transketolase. One answer is that transketolase gene *tktA* is downregulated approximately 10-fold during the later stages of the fermentation as supported by the microarray data (Table 16). However, the specific activities of transketolase remained the same over the course of the fermentation run of *E. coli* KL3/pJY1.216A (approximately 0.02 U/mg, data not shown). Another answer is that accumulation of superoxide (oxidative stress) during the fermentation process inhibits the synthesis of E4P by blocking an intermediate in the transketolase reaction.[20] The second answer is supported by two evidences. First, the upregulation of superoxide removing genes (*sodA* and *katE*, Table 20) and E4P producing genes (*talA* and *talB*, Table 16) in high yield cells suggests that E4P production correlates with superoxide removing. Second, since 3-dehydroshikimic acid as an antioxidant can react with $O_2$ and $H_2O_2$ to form gallic acid both in vitro and in vivo (mechanism is still unknown), gallic acid accumulation in the culture medium by the DHS producer can be viewed as an indication of high oxidative stress inside the cells. Once gallic acid was accumulating significantly in the culture medium during the later stages of the fermentation, the DHS biosynthesis ceased or stopped.

To further increase the availability of E4P, two options are available: increase the expression level of *tktA* gene by replacing the native *tktA* promoter with an enhanced expression promoter such as *aroF* promoter in a new plasmid similar to pJY1.216A; insert the superoxide removing gene *sodA* into pJY1.216A to construct another new

plasmid. These two plasmids will be transformed to *E. coli* KL3 and tested under fed-batch fermentation conditions. The attendant increases in the concentrations of synthesized DHS would be interpreted as indicating that carbon flow directed into the shikimate pathway is limited by the availability of E4P.

## REFERENCE


1 (a) Yi, J.; Li, K.; Draths, K. M.; Frost, J. W. Modulation of PEP synthase Expression Increases Shikimate Pathway Product Yields in *E. coli*. *Biotechnol. Prog.* **2002**, *18*, 1141-1148. (b) Chandran, S. S.; Yi, J.; Draths, K. M.; von Daeniken, R.; Weber, W.; Frost, J. W. Phosphoenolpyruvic Acid Availability and the Biosynthesis of Shikimic Acid. *Biotechnol. Prog.* **2003**, *19*, 808-814. (c) Yi, J.; Draths, K. M.; Li, K.; Frost, J. W. Altered Glucose Transport and Shikimate Pathway Product Yields in *E. coli*. *Biotechnol. Prog.* **2003**, *19*, 1450-1459.

2 (a) Gonzalez, R.; Tao, H.; Shanmugam, K. T.; York, S. W.; Ingram, I. O. Global gene expression differences associated with changes in glycolytic flux and growth rate in *Escherichia coli* during the fermentation of glucose and xylose. *Biotechnol. Prog.* **2002**, *18*, 6-20. (b) Oh, M. K.; Rohlin, L.; Kao, K. C.; Liao, J. C. Global expression profiling of acetate-grown *Escherichia coli*. *J. Biol. Chem.* **2002**, *277*, 13175-13183. (c) Oh, M. K.; Liao, J. C. Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*. *Biotechnol. Prog.* **2000**, *16*, 278-286.

3 (a) Yoon, S. H.; Han, M. –J.; Lee, S. Y.; Jeong, K. J.; Yoo, J. –S. Combined Transcriptome and Proteome Analysis of *Escherichia coli* During High Cell Density Culture. *Biotechnol. Bioeng.* **2003**, *81*, 753-766. (b) Khodursky, A. B.; Peter, B. J.; Cozzarelli, N. R.; Botstein, D.; Botstein, P. O.; Yanofsky, C. DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12170-12175.

4 (a) Tao, H.; Gonzalez, R.; Martinez, A.; Rodriguez, M.; Ingram, L. O.; Preston, J. F.; Shanmugam, K. T. Engineering a homo-ethanol pathway in *Escherichia coli*: Increased glycolytic flux and levels of expression of glycolytic genes during xylose fermentation. *J. Bacteriol.* **2001**, *183*, 2979-2988. (b) Tseng, G. C.; Oh, M. K.; Rohlin, L.; Liao, J. C.; Wong, W. H. Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variation and assessment of gene effects. *Nucleic Acids Res.* **2001**, *29*, 2549-2557.

5 Arfin, S. M.; Long, A. D.; Ito, E. T.; Tolleri, L.; Richle, M. M.; Paegle, E. S.; Hatfield, G. W. Global gene expression profiling in *Escherichia coli* K12: the effect of integration host factor. *J. Biol. Chem.* **2000**, *275*, 29672-29684.

6 Samules, M. L.; Witmer, J. A. *Statistics for the Life Sciences*. Prentice Hall, New Jersey, 1999.

7 Hartigan, J. *Clustering Algorithms*. New York: John Wiley & Sons, 1975.

8 Eisen, M. B.; Spellman, P. T.; Brown, P. O.; Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14863-14868.

9 Webb, C.; Scollon, S.; Miller, J.; Teh, B. T. Gene expression profiling of endocrine tumors by microarray analysis. *Current Opinion in Endocrinology & Diabetes* **2003,** *10*, 162-167.

10 *Biochemical pathways: an atlas of biochemistry and molecular biology*; Gerhard M. , Eds.; Wiley; Heidelberg; Spektrum; New York, 1999.

11 Karp, P. D.; Riley, M.; Saier, M.; Paulsen, l. T.; Collado-Vides, J.; Paley, S. M.; Pellegrini-Toole, A.; Bonavides, C.; Gama-Castro. The EcoCyc Database. *Nucleic Acids Res.* **2002**, *30*, 56-58.

12 Minoru, K. The KEGG database. NOVARTIS FOUNDATION SYMPOSIUM **2002**, *247*, 91-101; discussion 101-3, 119-28, 244-52.

13 Overbeek, R.; Larsen, N.; Pusch, G. D.; Dsouza, M.; Selkov, E.; Kyrpides, N.; Fonstein, M.; Maltsev, N.; SELKOV, E. WIT: integrated system for high-throughput genome sequence analysis and metabolic reconstruction. *Nucleic Acids Res.* **2000**, *28*, 123-125.

14 Fraenkel, D. G. Glycolysis. In *Escherichia coli and Salmonella*; Neidhardt, F. C., Ed.; ASM press: Washington, D. C., 1996; pp 189-198.

15 (a) Koebmann, B. J.; Westerhoff, H. V.; Snoep, J. L.; Nilsson, D.; Jensen, P. R. The Glycolytic Flux in *Escherichia coli* Is Controlled by the Demand for ATP. *J. Bacteriol.* **2002**, *184*, 3909-3916. (b) Koebmann, B. J.; Westerhoff, H. V.; Snoep, J. L.; Solem, C.; Pedersen, M. B.; Nilsson, D.; Michelsen, O.; Jensen, P. R. The extent to which ATP demand controls the glycolytic flux depends strongly on the organism and conditions for growth. *Mol. Biol. Reports* **2002**, *29*, 41-45.

16 Causey, T. B.; Zhou, S.; Shanmugam, K. T.; Ingram, L. O. Engineering the metabolism of *Escherichia coli* W3110 for the conversion of sugar to redox-neutral and oxidized products: Homoacetate production. *Proc. Nat. Acad. Sci.* **2003**, *100*, 825-832.

17 Pearse, A. J.; Wolf, R. E. Determination of the growth rate-regulated steps in expression of the *Escherichia coli* K-12 *gnd* gene. *J.Bacteriol.* **1994**, *176*, 115-122.

18 Lu, J.-L.; Liao, J. C. Metabolic engineering and control analysis for production of aromatics: Role of transaldolase. *Biotechnol. Bioeng.* **1997**, *53*, 132-138.

19 Cronan, J. E.; Laporte, D. Tricarboxylic acid cycle and glyoxylate bypass. In *Escherichia coli and Salmonella*; Neidhardt, F. C. Ed.; ASM Press: Washington, D. C., 1999; pp206-216.

20 Benov, L.; Fridovich, I. Why Superoxide Imposes an Aromatic Amino Acid Auxotrophy on *Escherichia coli*. *J. Biol. Chem.* **1999**, *274*, 4202-4206.

21 (a) Asami, S.; Akazawa, T. Enzyme Formation of Glycolate to Chromatium: Role of Superoxide Radical in Transketolase-type Mechanism. *Biochemistry*, **1977**, *16*, 2201-2207. (b) Takabe, T.; Asami, S.; Akazawa, T. Glycolate Formation Catalyzed by Spinach Leaf Transketolase Utilizing the Superoxide Radical. *Biochemistry*, **1980**, *19*, 3985-3989.

22 Epstein, W.; Wieczorek, L.; Siebers, A.; Altendorf, K. Potassium Transport in *E. coli*: genetic and biochemical characterization of the potassium-transporting ATPase. *Biochem. Soc. Trans.* **1984**, *12*, 235-236.

23 Foster, J. W. Microbial responses to acid stress. In *Bacterial Stress Responses*; Storz, G., Hengge-Aronis, R., Eds.; ASM Press: Washington, 2000; pp 99-115.

24 Almiron, M.; Link, A. J.; Furlong, D.; Kolter, R. A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev.* **1992**, *6*, 2646-2654.

25 Dukan, S.; Touati, D. Hypochlorous acid stress in *Escherichia coli*: resistance, DNA damage, and comparison with hydrogen peroxide stess. *J. Bacteriol.* **1996**, *178*, 6145-6150.

26 (a) Ogino, T.; Garner, C.; markley, J. L.; Herrmann, K. M. "Biosynthesis of aromatic compounds:$^{13}$C NMR spectroscopy of whole *Escherichia coli* cells. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 5828-5832. (b) Weaver, L. M.; Herrmann, K. M. Cloning of an AroF Allele Encoding a Tyrosine-Insensitive 3-Deoxy-D-*aroabino*-heptulosonate 7-phosphate synthase. *J. Bacteriol.* **1990**, *172*, 6581-6584.

27 Mikola, M. R.; Widman, M. T.; Worden, R. M. In situ mutagenesis and chemotactic selection of microorganisms in a diffusion gradient chamber. *Appl. Biochem. Biotechnol.* **1998**, *70-72*, 905-918.

[28] Li, K.; Mikola, M. R.; Draths, K. M.; Worden, R. M.; Frost, J. W. Fed-Batch Fermentor Synthesis of DHS Using Recombinant *Escherichia coli*. *Biotechnol. Bioeng.* **1999**, *64*, 61-73.

29 (a) DeLucia, A.; Schoner, R.; Herrmann, K. M. Selective Isoenzyme Inactivation in Biosynthesis of Aromatic Compounds. *Abstracts of Papers*, Xith International Congress of Biochemistry, Toronto, Canada, 1979, 04-2-S4. (b) Tribe, D. E.; Pittard, J. Hyperproduction of Tryptophan by *Escherichia coli*: Genetic Manipulation of the Pathways Leading to Tryptophan Formation. *Appl. Environ. Microbiol.* **1979**, *38*, 181-190.

30 Cadwell, R. C.; Joyce, G. F. Mutagenic PCR. *PCR Methods Applic.* **1994**, *4*, S136-S139.

31 Akowski, J. P.; Bauerle, R. Steady-State Kinetics and Inhibitor Binding of 3-Deoxy-D-*aroabino*-heptulosonate-7-Phosphate Synthase (Tryptophan sensitive) from *Escherichia coli. Biochemistry* **1997**, *36,* 15817-15822.

32 Yoshimi, K.; Kumiko, T.; Osamu, K. Mutational Analysis of the Feedback Sites of Phenylalanine-Sensitive 3-Deoxy-D-*arabino*-Heptulosonate- 7-Phosphate Synthase of *Escherichia coli. Appl. Environ. Microbiol.* **1997**, *63*, 761-762.

33 (a) Kikuchi, T.; Sotochi, N.; Fukase, K.; Kojima, H.; Kurahashi, O.; matsui, Y. Aromatic amino acid manufacture enhancement with recombinant microorganism. JP Patent 04248983, 1993. (b) Tonouchi, N.; Kojima, H.; Matsui, H. The use of feedback-insensitive enzymes in the production of aromatic amino acids by fermentation. Eur. Pat. Appl. 488424, 1992.

34 Mikola, M. R.; Widman, M. T.; Worden, R. M. In situ mutagenesis and chemotactic selection of microorganisms in a diffusion gradient chamber. *Appl. Biochem. Biotechnol.* **1998**, *70-72*, 905-918.

35 Walker, N. J. A Technique Whose Time Has Come. *Science* **2002**, *296*, 557-558.

# CHAPTER 4

## Studies of shikimate export

### Background

As is evident from current genome analyses, a substantial number of bacterial genes encode membrane transport proteins. These proteins enable the controlled exchange of molecules between the cell and its environment. In *Escherichia coli*, more than 400 genes encode transport proteins based on sequence homology to known transport genes.[1] However, a lot of these putative transport proteins are functionally undefined. Many characterized transport proteins are used to import nutrients such as carbohydrates, ions, amino acids, or peptides. Other proteins are known to act as exporters. Most export proteins catalyze the export of noxious substances. Examples are multidrug resistance proteins,[2] metal resistance proteins,[3] and polysaccharide export proteins.[4] Recent studies have shown that there are also proteins that catalyze the export of sugars[5] and amino acids.[6] The first export protein to be identified was LysE[6c] of *Corynebacterium glutamicum.* Exression of LysE is necessary during growth on complex medium or in the presence of peptides rich in L-lysine or L-arginine.[7] Under such special growth conditions, L-lysine or L-arginine might accumulate to toxic levels, a situation which is prevented by their export. Both L-lysine and L-arginine are exported by LysE at a rate of 0.75 nmol $min^{-1}$ (mg dry wt) $^{-1}$.[7] Proteins homologous to LysE are widespread and occur in bacteria and archaea.[8]

Studies on the peptide uptake systems of *Staphylococcus aureus,*[9] *Streptococcus faealisi,* and *E. coli*[10] indicated that amino acids are exported that was components of peptides initially consumed by these bacteria. The export of amino acids also occurred with *Lactococcus lactis* grown on milk[11] or in the presence of milk-derived peptides.[12] Those phenomena suggested the existence of more exporters of amino acids. Examination of the export of L-threonine from *E. coli* and the export of L-glutamate, L-lysine, L-isoleucine, and

L-threonine from *Corynebacterium glutamicum* suggests the existence of active transporters.[13] Based on these studies, it recently became possible to identify specific export carriers at the molecular level. This has led to the identification of novel transport families and to insights into how bacteria control the intracellular concentrations of metabolites. Some examples of identified export proteins for amino acids and sugars include: (a) the ThrE[14] protein of *Corynebacterium glutamicum*, which is involved in threonine export; (b) the RhtB[15] protein of *E. coli*, which exports the homoserine lactones, threonine and other amino acids; (c) the Orf299[6b] protein of *E. coli*, which is involved in cysteine export; (d) the YedA[5a] protein of *E. coli*, which exports L-arabinose; and (e) the SetA[5b] protein of *E. coli* for the export of glucose.

The export protein serves as a valve that exports a metabolite that increases to excessive intracellular concentrations. If the intracellular concentration of a microbe-synthesized product exceeds the extracellular concentration, the rate of export may limit product concentrations and yields. In the literature, the export of L-lysine,[16] L-threonine[14] and L-isoleucine[17] in *Corynebacterium glutamicum* was an impediment to the development of strains that synthesized these amino acids. One strategy to achieve high-yield production in such fermentation process is to increase the rate of export of the product amino acid. For example, threonine production by *Escherichia coli* was enhanced by overexpression of the *E. coli rhtB* and *rhtC* genes or by heterologous overexpression of the *thrE* gene encoding the *Corynebacterium glutamicum* threonine exporter.[18] Shikimic acid and 3-dehydroshikimic acid are hydrophilic molecules and are likely to export out of the cells by a carrier-mediated process instead of diffusion. In this chapter active export of shikimic acid was implicated, the hypothesis that the rate of export constitutes a bottleneck during the fermentation was tested, and two approaches to identify the carrier were explored.

**Shikimate export is a carrier-mediated process and not a bottleneck for shikimate production.**

The functional analysis of export processes is more complicated than studies of nutrient uptake. The use of labeled transport substrates, which is the method of choice for kinetic analysis of uptake processes is restricted since most substrates are metabolized in the cell. In general, analysis of export systems requires quantification of intracellular and extracellular concentrations of the product or biosynthetic intermediate. In order to establish intracellular concentrations, the correct cytoplasmic volume must be determined. In general, this is carried out by using silicone oil centrifugation to separate the cell from the culture medium. The cytoplasmic volume has been measured by double labeling techniques and a cytoplasmic volume of 1.7 μL/ mg dry cell weight was used in our calculation.[19] Because of the highly unfavorable volume ratio (internal volume<< external volume), the exact quantification of intracellular concentration is difficult. A recently reported chromogenic assay facilitated rapid and easy quantification of shikimic acid concentration.[20]

*E. coli* shikimate producers SP1.1/pKD12.138 and SP1.1/pKD15.071 were used to examine shikimate export system by determining the intracellular and extracellular shikimate concentrations during the fermentation.[21] The host strain SP1.1 used for synthesis of shikimic acid was derived from *E. coli* RB791 by homologous recombination of the *aroB* gene into the *serA* locus followed by successive P1 phage-mediated transductions to introduce *aroL*478*::Tn10* and *aroK::Cm*$^R$ mutations. The resulting RB791 *serA::aroB aroL::Tn10 aroK::Cm*$^R$ lacked 3-phosphoglycerate dehydrogenase (SerA) activity as well as the activities of both isozymes of shikimate kinase (AroK and AroL). In the absence of shikimate kinase, carbon flow directed into the common pathway was unable to proceed beyond biosynthesis of shikimic acid, which accumulates extracellularly in the culture medium. Plasmid pKD12.138 contained *aroF*$^{FBR}$, *tktA*, *serA*, $P_{tac}$*aroE* while pKD15.071 contained a *ppsA* gene in addition to all of the aforementioned inserts in pKD12.138.

At timed intervals during the fermentation, an aliquot of *E. coli* cells was centrifuged and the resulting cell pellet was resuspended in ice-cold buffer. In order to measure intracellular concentrations, the cellular reactions have to be rapidly terminated by mixing with a quenching agent such as perchloric acid. In the silicone oil centrifugation method,[22] the cells are on top, the silicon oil layer is in the middle and the acid layer is at the bottom prior to centrifugation. During centrifugation, the cells pass through the silicone oil layer and are stripped of their surrounding medium. The cell pellets in the perchloric layer were then disrupted by sonication, and the mixture was neutralized. After centrifugation, the shikimic acid levels in the supernatant were determined using a chromogenic assay.[20] In this assay, shikimic acid was oxidized by periodic acid to give *trans*-aconitic acid and a dialdehyde and then basicified to pH 10 with addition of NaOH to give an intense yellow characteristic of dialdehyde formation. Only shikimic acid, quinic acid and tryptophan have been found to produce a chromophore having a yellow color after treatment with periodic acid and subsequent basicification.

**Table 24. Shikimate intracellular and extracellular concentration during the fermentation of SP1.1/pKD12.138 under glucose-rich conditions.**

| Time (h) | $OD_{600}$ | $SA_{Intra}$ (mM) | $SA_{Extra}$ (mM) |
|---|---|---|---|
| 12 | 7.5 | 4 | 3 |
| 18 | 57 | 10 | 19 |
| 24 | 54 | 10 | 95 |
| 30 | 50 | 6 | 145 |
| 36 | 49 | 10 | 185 |
| 42 | 47 | 8 | 225 |
| 48 | 49 | 10 | 231 |
| 54 | 47 | 8 | 278 |
| 60 | 45 | 7 | 305 |

**Table 25. Shikimate intracellular and extracellular concentration during the fermentation of SP1.1/pKD15.071 under glucose-rich conditions.**

| Time (h) | $OD_{600}$ | $SA_{Intra}$ (mM) | $SA_{Extra}$ (mM) |
|---|---|---|---|
| 12 | 4.1 | 10 | 6 |
| 18 | 24 | 9 | 19 |
| 24 | 52 | 28 | 82 |
| 30 | 49 | 11 | 171 |
| 36 | 49 | 10 | 215 |
| 42 | 49 | 8 | 240 |
| 48 | 42 | 19 | 300 |
| 54 | 44 | 9 | 329 |
| 60 | 40 | 6 | 387 |
| 66 | 42 | 6 | 389 |
| 72 | 42 | 7 | 351 |

The intracellular shikimate concentration of 7-30 mM measured over the course of fermentation runs of *E. coli* SP1.1/pKD12.138 and SP1.1/pKD15.071 was low compared to the extracellular concentration of shikimate, which reached a maximum concentration of approximately 400 mM (Tables 24 and 25). In the literature, Morbach et al. measured intracellular and extracellular concentration of L-isoleucine during cultivation of *Corynebacterium glutamicum* SM13. The intracellular L-isoleucine concentration was always higher relative to the extracellular concentration of L-isoleucine. The intracellular concentration rose to 110 mM, whereas extracellularly the extracellular concentration of L-isoleucine measured to only a 60 mM concentration.[17] Similar analysis of the *E. coli* L-threonine producer strain BKIIM B-3396/pVIC40 also revealed that the intracellular L-threonine concentration exceeded the extracellular concentration over the course of the entire fermentor run.[18a] At times, the 250 mM intracellular L-threonine concentration was 10 times

higher than the extracellular concentration. The low intracellular shikimate concentration during its microbial synthesis under fermentor conditions has two implications. First, the export of shikimate is not likely to be the rate-limiting step for the production of shikimic acid since no accumulation of intracellular shikimic acid was observed. Second, because the export of shikimic acid is not driven by the difference between the intracellular and extracellular concentration, the export of shikimic acid is most likely a carrier-mediated process.

## Approaches to identify the loci involved in the export of shikimate

### First approach: isolation of export-defective mutants by mini-Tn10 mutagenesis

The same export protein may exist for both shikimate and DHS because of their structure similarity. The first approach for identification of loci involved in the export of shikimate and DHS used mini-Tn10 mutagenesis of the shikimate-producing strain *E. coli* SP1.4 (*serA*::*aroF*$^{FBR}$*aroB aroL*::*Tn10*)[23] to disrupt loci involved in shikimate export. Carbon flow directed into the shikimate pathway is increased in *E. coli* SP1.4 with the genomic insertion of *aroF*$^{FBR}$ into the *serA* locus. By inactivating only one of the two loci encoding shikimate kinase with the *aroL*::*Tn10* gene disruption, *E. coli* SP1.4 accumulated 0.2 mM DHS and 0.3 mM shikimic acid in its culture medium under shake flask conditions (see experimental) while allowing enough carbon flow through the shikimate pathway to provide all of its aromatic amino acid and aromatic vitamin requirements. Phenylalanine (0.6 mM) and tyrosine (0.9 mM) were also accumulated in the culture medium by *E. coli* SP1.4 while no tryptophan was accumulated in the culture medium. *E. coli* SP1.4 was submitted to transposon mutagenesis[24] using λNK1324 obtained from ATCC, which is a mini-Tn10 construct marked with an insert encoding for resistance to chloramphenicol. Infection of *E. coli* SP1.4 with λNK1324 was followed by selection for resistance to chloramphenicol. These chloramphenicol-resistant colonies were then plated on solid

medium containing *E. coli* AB2847, which is an *aroB* mutant that requires aromatic supplementation. Halos of *E. coli* AB2847 grew around *E. coli* SP1.4 mutants that retained the ability to export DHS and shikimic acid. In contrast, no halo of *E. coli* SP1.1 (*aroKaroL* mutant that requires aromatic supplementation) was observed around *E. coli* SP1.4. Since the difference between *E. coli* SP1.1 and *E. coli* AB2847 is that *E. coli* SP1.1 cannot utilize DHS and shikimic acid for aromatics biosynthesis, DHS and shikimic acid exported by *E. coli* SP1.4 was metabolized by *E. coli* AB2847 to provide the tryptophan and aromatic vitamins requirement for this auxotroph. *E. coli* SP1.4 mutants that lost export capability were among those lacking halos or those displaying reduced halo sizes.

For isolation of transposition events into the bacterial chromosome, bacteriophages are the most convenient type of delivery vehicle. Phage λNK1324 carrying the transponson mini-Tn10 (approximately 1.4 kb) can be introduced into the host cell *E. coli* SP1.4 under conditions where the phage genome didn't replicate, kill or stably integrates into the host cell.[24] Mini-Tn10 was chosen so as to ensure random, stable insertions into the *E. coli* SP1.4 chromosome. Transposons that do not contain a transposase gene within their boundaries are generally refered to as minitransposons. By using the minitransposons such as mini-Tn10, stable transposon insertions that are unable to undergo additional rounds of transposition can be obtained. Another advantage of λNK1324 is its usage of mutant transposases that exhibit a much lower degree of insertion specificity than wild-type.[24] The λNK1324 phage lysate was prepared using standard procedures and then used to infect *E. coli* SP1.4 under conditions that preclude lysogen formation.[25] The cells that have acquired the chloramphenicol-resistance marker from the mini-Tn10 were selected on rich plates containing chloramphenicol. Normally 100 to 200 colonies per plate were obtained by optimizing the ratio of phage to cells. Serial dilutions ("pies") on rich plates containing chloramphenicol were used to isolate single colonies. The single chloramphenicol-resistant colonies were replicate plated on M9 salts solid medium containing L-phenylalanine, L-tyrosine and L-serine along with strain *E. coli* AB2847 at a density of $10^7$ cells per mL.

After 24 h of growth in the solid medium, colonies were inspected for reduced shikimic acid or DHS export as indicated by reduced halo formation of *E. coli* AB2847. With the supply of L-phenylalanine and L-tyrosine in the selection plate, *E. coli* SP1.4 mutants with the reduced export of L-phenylalanine and L-tyrosine showed normal halo size. From approximately 9000 isolated colonies, 16 colonies were selected that showed significantly decreased halo formation. Of the selected colonies 8 of 16 showed normal growth on M9 salts solid medium with L-serine indicating no essential gene in these colonies was disrupted by mini-Tn10 insertion mutagenesis. All the other 8 colonies were auxotrophic, suggesting a cross feeding between *E. coli* SP1.4 mutants and *E. coli* AB2847.

## Y-linker PCR technique to identify the sequences that flank the transposon insertions

The method of Kwon et al.[26] that requires the sequence information of only transposon-specific sequences (Figure 33) was used for specific amplification of DNA sequences that flank a transposon insertion. The genomic DNA isolated from a mini-Tn10 mutant was digested to completion with *Nla*III, a restriction enzyme with a 4-bp recognition site. The digested DNA was ligated to Y linker (Figure 33) designed to have a 3' overhang complementary to the sticky end generated by the *Nla*III (CATG). PCR was then performed using the Y-linker primer and mini-Tn10 specific primer. The Y-linker has a region of noncomplementary sequence on the 5' end. The Y-linker primer is made from this Y region and cannot anneal to the linker itself. Therefore, the fragments that do not have the mini-Tn10 sequence were not amplified in the PCR reaction. However, the mini-Tn10 primer anneals to the fragments containing mini-Tn10 sequences during the first PCR cycle and extends DNA synthesis into the Y region of the ligated Y linker, where the Y linker primer can anneal. Starting from the second cycle, the fragments containing mini-Tn10 sequences and some genomic sequences were selectively amplified. The function of

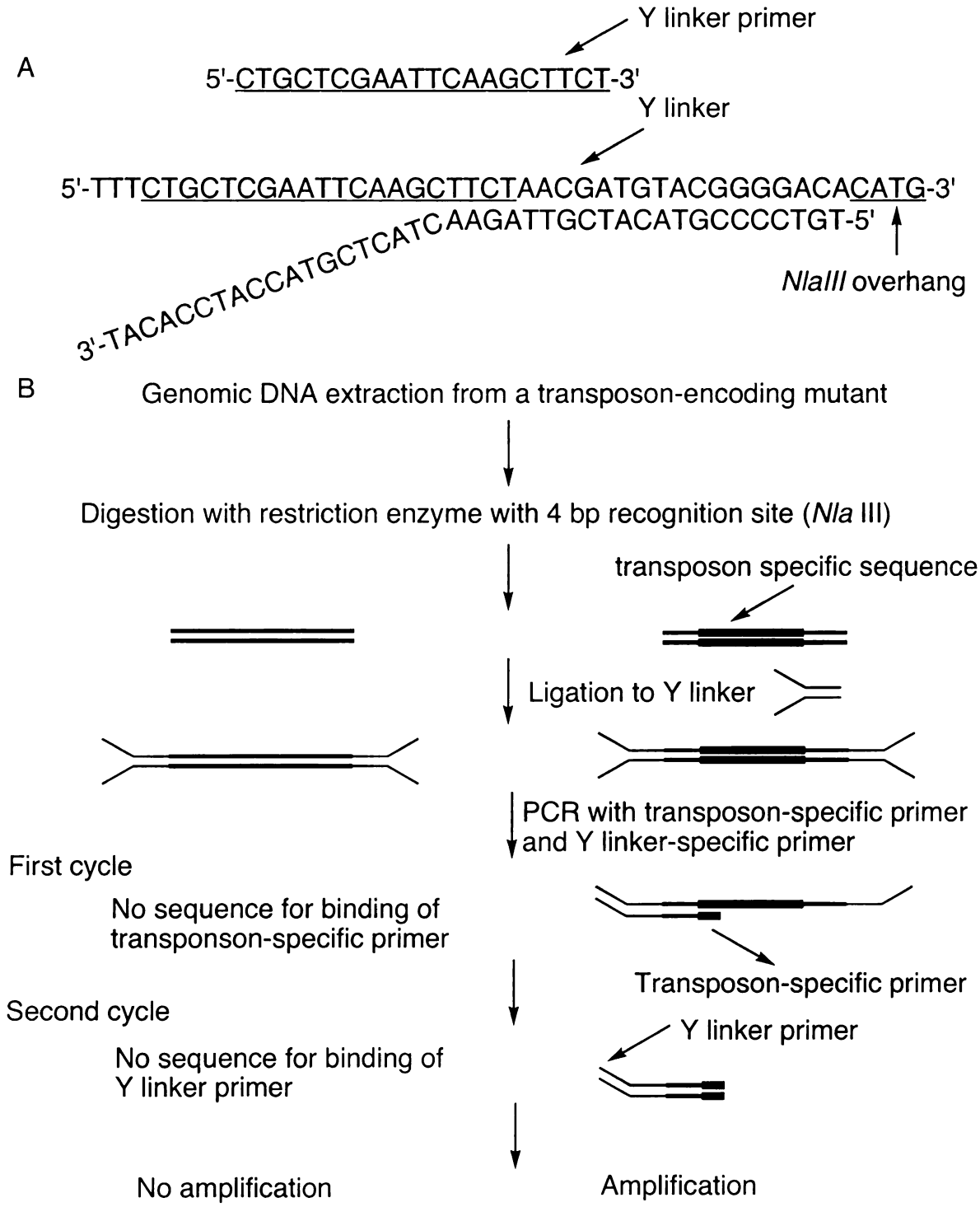



**Figure 33. (A) Y-linker. (B) PCR of transponson-flanking sequences.**

the noncomplementary region on the 5' end of the Y linker was to prevent nonspecific background amplification. If the noncomplementary region on the 5' end of the Y linker

was not present, the digested fragments that do not have a mini-Tn10 sequence also would provide an annealing site for Y linker primer and would be amplified with two Y linker primers annealed at both ends.

**Table 26. SP1.4 mutants.**

| Mutant ID | Growth on M9/glu/ser | Location of Tn10 insertion | Gene function |
|---|---|---|---|
| 1 | + | *ilvI* | Acetolactate synthase isozyme III large subunit |
| 2 | + | *ilvI* | Acetolactate synthase isozyme III large subunit |
| 3 | + | *narU* | Nitrite extrusion protein |
| 4 | + | *ilvI* | Acetolactate synthase isozyme III large subunit |
| 5 | + | *ydaQ* | Unidentified protein |
| 6 | + | *tdcA* | tdcABC Operon transcriptional activator |
| 7 | + | *ksgA* | S-adenosylmethionine-6-N', N'-adenosyl (rRNA) dimethyltransferase |
| 8 | + | *ykgH* | Unidentified protein |
| 9 | - | *argE* | Acetylornithine deacetylase |
| 10 | - | *argG* | Argininosuccinate synthase |
| 11 | - | *argH* | Argininosuccinate lyase |
| 12 | - | *cysN* | Sulfate adenylyltransferase |
| 13 | - | *purH* | Aminoimidazole carboxamide ribonucleotide transformylase / IMP cyclohydrolase |
| 14 | - | *metE* | Cobalamin-independent homocysteine Transmethylase |
| 15 | - | *cysN* | Sulfate adenylyltransferase |
| 16 | - | *pdxA* | Phospho-hydroxy-threonine dehydrogenase/decarboxylase |

As expected, a PCR product with length of 0.2-1.0 kb was amplified for each mini-Tn10 *E. coli* SP1.4 mutant. The PCR product was purified and was sequenced in the Genomics Technology Support Facility at Michigan State University. The resulting mini-Tn10 flanking sequences were compared in the National Center for Biotechnology Information (NCBI) database to determine the genomic location of the mini-Tn10 insertion.

The results of 16 mutants are summarized in Table 26. Mutants 9-16 included in Table 26 were auxotrophic for different nutrients: amino acids, nucleotides and cofactors. The mutated *argE*, *argG* and *argH* loci are involved in the biosynthesis of arginine. The employed selection protocol thus seemed to be a surprisingly good way of selecting mutations in genes encoding arginine biosynthetic proteins. One possible explanation is that an arginine residue plays an important role in the protein-mediated export of DHS or shikimic acid. If the supply of arginine is limited, the export machinery for DHS or shikimic acid slows down. Of the 16 mutants, only NarU is annoted as a transport protein (for nitrite export).[27]

To further characterize the 16 mutants, two methods were used. First, the intracellular and extracellular shikimate concentrations of the 16 mutants were measured under shake flask conditions. Secondly, the impact of the mutation under fed-batch fermentation conditions was evaluated. The mini-Tn10 insertion in the genome of SP1.4 mutants (gene knockout) was transferred to the DHS producer *E. coli* KL3 by P1 phage transduction. Shikimic acid producer *E. coli* SP1.1 was not suitable for P1 transduction because of the chloramphenicol marker in the genome. The *E. coli* KL3 transductants were evaluated for DHS production after transformation with plasmid pJY1.216A under glucose-rich fermentor conditions. The major difference between those two methods (shake flask conditions and fed-batch fementation conditions) is that a much higher concentration of DHS is accumulated extracellularly in the second method than the concentration of shikimic acid accumulated in the first method.

**Determination of intracellular shikimate concentrations of *E. coli* SP1.4 mutants under shake flask conditions**

*E. coli* SP1.4 mutants were grown in 100 mL LB medium for 12 h, washed with 100 mL M9 salts medium and resuspended in 100 mL M9 salts medium supplement with L-serine and glucose as the sole carbon source. After cultivation for 24 h at 37°C, the

intracellular and extracellular concentrations of shikimic acid were determined using silicon oil centrifugation and chromogenic assay as described in the previous section. Table 27 shows the results of the 16 mutants. Some *E. coli* SP1.4 mutants (*ilvI*, *ydaQ*, *tdcA*, *ksgA*, *argE*, *argH*, *pdxA*) retained a 2-3 fold more intracelluar shikimic acid concentration than the control *E. coli* SP1.4. However, accumulation of shikimic acid did not reach levels comparable to those reported during export of L-lysine,[16] L-isoleucine[17] and L-threonine.[14]

**Table 27. Intracellular and extracellular shikimic acid accumulation of SP1.4 mutants under shake flask conditions.**

| Mutant ID | Location of Tn10 insertion | $SA_{Intra}$ (mM) | $SA_{Extra}$ (mM) |
|---|---|---|---|
| SP1.4 | | 0.8 | 0.5 |
| 1 | *ilvI* | 2.4 | 0.4 |
| 2 | *ilvI* | 2.4 | 0.4 |
| 3 | *narU* | 1.1 | 0.4 |
| 4 | *ilvI* | 1.5 | 0.4 |
| 5 | *ydaQ* | 2.2 | 0.3 |
| 6 | *tdcA* | 1.8 | 0.4 |
| 7 | *ksgA* | 2.0 | 0.4 |
| 8 | *ykgH* | 1.2 | 0.3 |
| 9 | *argE* | 1.7 | 0.4 |
| 10 | *argG* | 1.4 | 0.4 |
| 11 | *argH* | 2.4 | 0.3 |
| 12 | *cysN* | 0.8 | 0.4 |
| 13 | *purH* | 1.0 | 0.4 |
| 14 | *metE* | 0.7 | 0.4 |
| 15 | *cysN* | 1.0 | 0.4 |
| 16 | *pdxA* | 1.6 | 0.5 |

**Table 28. Production of DHS by KL3 derivatives/pJY1.216A under glucose-rich conditions.**

| Construct | DHS Titer (g/L) | DHS Yield (%) | Total Yield (%) |
|---|---|---|---|
| KL3/pJY1.216A | 69 | 35 | 51 |
| KL3*tdcA*/pJY1.216A | 71 | 38 | 53 |
| KL3*ilvI*/pJY1.216A | 52 | 28 | 37 |
| KL3*narU*/pJY1.216A | 65 | 35 | 49 |
| KL3*ydaQ*/pJY1.216A | 69 | 36 | 51 |
| KL3*ksgA*/pJY1.216A | 46 | 34 | 43 |
| KL3*ykgH*/pJY1.216A | 61 | 35 | 49 |

**DHS titers and yields by KL3 derivatives**

The specific gene knockouts in *E. coli* KL3 were obtained by phage P1 transduction using *E. coli* SP1.4 mutants1-8 (see Table 26) as the donor strain. The *E. coli* KL3 derivatives with specific gene knockouts were transformed with plasmid pJY1.216A and tested for the production of DHS under glucose-rich fed-batch fermentor conditions with 0.5 mL 100 mM IPTG added as described previously.[28]

The results (Table 28) indicated that except *E. coli* KL3*ilvI* and KL3*ksgA*, other KL3 mutants produced almost the same amount of DHS as the control strain *E. coli* KL3/pJY1.216A. *E. coli* KL3*ksgA*/pJY1.216A grew poorly and the maximum dry cell weight during the course of the fermentation run of *E. coli* KL3*ksgA*/pJY1.216A was only two thirds of the maximum dry cell weight observed during the course of the fermentation of *E. coli* KL3/pJY1.216A (data not shown). *E. coli* KL3*ilvI*/pJY1.216A grew normally and produced significantly less DHS with lower yield relative to the control strain *E. coli* KL3/pJY1.216A (Table 28). However, the reason for reduced DHS production by *E. coli* KL3*ilvI*/pJY1.216A is probably due to reduction of PEP synthase activity rather than the disruption of shikimate export machinery. The PEP synthase activities for *E. coli* KL3*ilvI*/pJY1.216A were determined to be much less throughout the fermentation relative to

the control fermentation of *E. coli* KL3/pJY1.216A (Table 29). To prove this hypothesis, plasmid pKL5.17A that didn't overexpress the PEP synthase was transformed to KL3*ilvI* and this time the *ilvI* gene knockout had no effect on the production of DHS relative to the control strain *E. coli* KL3/pKL5.17A (data not shown).

**Table 29. PEP synthase specific activities for fed-batch fermentation of KL3*ilvI*/pJY1.216A and KL3/pJY1.216A with 0.5 mL 100 mM IPTG addition under glucose-rich conditions.**

| Construct | PEP synthase specific activity (U/mg) | | |
|---|---|---|---|
| | 24 h | 36 h | 42 h |
| KL3*ilvI*/pJY1.216A | 0.02 | 0.02 | 0.02 |
| KL3/pJY1.216A | 0.07 | 0.06 | 0.05 |

The above results raised the hypothesis that there might exist multiple export carriers for shikimic acid (NarU may be one of them). The existence of multiple export carriers is reasonable because in contrast to the export of toxic compounds and of fermentative end products, there seems to be no need for the evolution of specific systems for export of hydroaromatics under normal physiological conditions.

**Second approach: resistant to fluoroshikimic acid**

In the second approach, a plasmid library of *E. coli* genomic DNA was constructed and probed for enhanced resistance towards the antibiotic (*6S*)-6-fluoroshikimic acid.[29] The basis of this selection is the hypothesis that the shikimic acid export system can also export the structurally similar molecule (*6S*)-6-fluoroshikimic acid. The shikimate transport system (ShiA) has been shown to also transport (*6S*)-6-fluoroshikimic acid.[30] The overexpression of a gene encoding the shikimic acid export system in a plasmid should increase the rate of (*6S*)-6-fluoroshikimic acid export from the *E. coli* cytoplasm. Increased export of (*6S*)-6-fluoroshikimic acid should result in growth in the presence of concentrations of this antibiotic that completely prevent growth of wild-type *E. coli*. (*6S*)-6-

Fluoroshikimic acid was synthesized from shikimic acid by a modification[31] of literature procedures.[32]

**Table 30. Identified genes to resist (*6S*)-6-fluoroshikimic acid toxicity.**

| ID | Responsible genes in plamsid | FSA in the plate | | |
|---|---|---|---|---|
| | | 2 μg/ mL | 10 μg/ mL | 50 μg/ mL |
| 1 | *yibA, yibJ* | + | - | - |
| 2 | *b1497, yddA* | + | - | - |
| 3 | *entC* | + | + | + |
| 4 | *ampC, frdBCD* | + | + | - |
| 5 | *b2432, b2433, b2434* | + | - | - |
| 6 | *b2420, cysM* | + | - | - |
| 7 | *entC* | + | + | + |
| 8 | *pepA, holC* | + | - | - |
| 9 | *yidS* | + | - | - |
| 10 | *b2353* | + | - | - |
| 11 | *holC* | + | - | - |
| 12 | *entC* | + | + | + |
| 13 | *yfiP, yfiQ* | + | - | - |
| 14 | *fadL* | + | + | - |
| 15 | *glvB, glvC* | + | - | - |
| 16 | *yifB, ilvL* | + | - | - |
| 17 | *ppdABC* | + | - | - |
| 18 | *entC* | + | + | + |
| 19 | *yafO, yafN, dinP* | + | - | - |
| 20 | None | - | - | - |

A genomic library prepared from strain RB791 was constructed as follows. Chromosomal DNA was purified from a cell lysate.[33] The purified DNA was partially digested with *Sau*3A, and DNA fragments estimated to range in size from 3 to 10 kb were isolated from an agarose gel. The resulting DNA was inserted into pBR322, which has been digested with *Bam*HI and dephosphorylated with calf intestinal alkaline phosphatase.

The resulting ligation mixture was introduced by transformation into XL10-Gold ultracompetent cells from Stratagene, and the resulting transformants were selected at 37 °C on M9 minimal medium plates containing ampicillin and fluoroshikimic acid (0.5 μg/ mL). The 0.5 μg/ mL fluoroshikimic acid is adequate to prevent the growth of wild-type XL10-Gold cells. The plasmids that suppressed fluoroshikimic acid toxicity were isolated from 19 colonies that grew on the selection plate. The colonies were also replicate plated on solid medium with increasing concentrations (2 μg/ mL, 10 μg/ mL and 50 μg/ mL) of fluoroshikimic acid (Table 30).

The plasmids that suppressed fluoroshikimic acid toxicity were sequenced in the Genomics Technology Support Facility in Michigan State University and compared in the National Center for Biotechnology Information (NCBI) database to determine the responsible genes in the plasmid (Table 30). The only gene leading to resistance to 50 μg/mL fluoroshikimic acid is *entC* which encodes isochorismate synthase. In addition to *entC*, overexpression of *fadL*, which encodes the transport protein for long-chain fatty acids, or *ampC/frdBCD*, which encodes β-lactamase resistance and fumarate reductase in the host strain led to resistance towards 10 μg/mL fluoroshikimic acid. There were a number of genes leading to resistance towards 2 μg/mL fluoroshikimic acid. Some of those genes encode unidentified proteins, but none of them are transport proteins based on sequence homology. Some genes encode functional proteins such as DNA polymerase III (*holC),* damage inducible protein (*dinP*), PTS transport protein (*glvBC*), or proteins involved in the amino acid biosynthesis (*ilvL* and *cysM*). The difficulty of identifying a shikimate export protein by the selection of fluoroshikimic acid resistance is probably because of the low accumulation of fluoroshikimic acid inside the cells. If the export machinery inside the cells is already functioning well, increased expression will not likely lead to a selectable phenotype.

Fluoroshikimic acid appears to enter the cytoplasm of *E. coli* via the transport

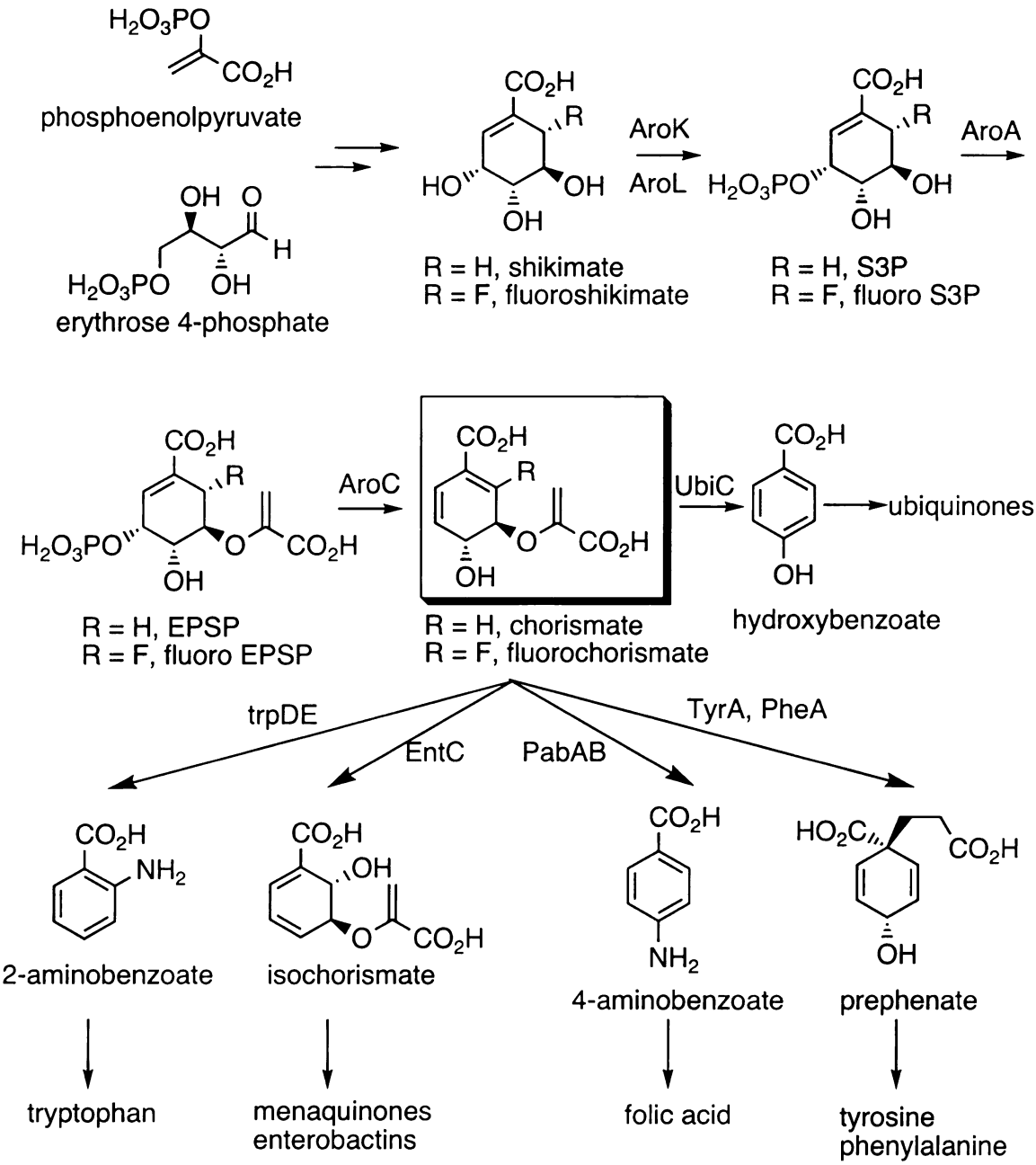



**Figure 34. Summary of the aromatic biosynthetic pathway.** Genetic loci are as follows: AroK, shikimate kinase; AroL, shikimate kinase; AroA, EPSP synthase; AroC, chorismate synthase; EntC, isochorismate synthase; PabAB, PABA synthase; TyrA, chorismate mutase; PheA, chorismate mutase; TrpDE, anthranilate synthase; UbiC, chorismate-pyruvate lyase.

system (ShiA) for shikimate.[30] It then becomes a substrate for shikimate kinase (AroK, AroL) and 5-enol-pyruvyl-shikimate-3-phosphate synthase (AroA) and chorismate

synthase[34] (AroC) thereby being transformed to 6-fluorochorismate (see Figure 34).[35] Thereafter, although the individual steps have not been elucidated, it has been suggested that 6-fluorochorismate inhibits the synthesis of 4-aminobenzoate (Figure 34), which is a key intermediate in the production of the folate coenzymes for one-carbon metabolism.[29] Davies et al. observed that addition of 4-aminobenzoic acid was sufficient to overcome the growth inhibition of fluoroshikimic acid.[29] Addition of aromatic amino acids did not overcome fluoroshikimic acid growth inhibition. In our study, the overexpression of *entC* gene encoding isochorismate synthase led to resistance to the highest concentration of fluoroshikimic acid (50 μg/mL). Our result suggests that the 6-flurochorimsate inhibited the isochorismate synthase responsible for the production of menaquinones and enterobactins. Reconciling this observation with the previously reported ability of 4-aminobenzoic acid supplementation to reverse fluoroshikimate growth inhibition requires further study.

## Conclusion

Two approaches to identify the shikimate export protein were attempted without success. The first approach attempted to identify a mutant deficient in shikimate export while the second approach was based on the overexpression of a shikimate export protein. The failure of both approaches is consistent with the possibility that multiple proteins exist for the export of shikimic acid. Thus the first approach failed to identify the shikimate export protein because mini-Tn10 mutagenesis can only knock out one gene at a time and the second approach failed because the intracellular concentration of (*6S*)-6-fluoroshikimic acid that accumulated is below the $K_m$ of the exporter. Furthermore, the determination that intracellular shikimic acid concentrations are much lower than extracellular concentrations suggests that export of this hydroaromatic is not a limiting factor during microbial synthesis.

## REFERENCE


1 Blattner, F. R.; Plunkett, G.; Bloch, C. A.; Perna, N. T.; Burland, V.; Biley, M.; et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **1997**, *277*, 1453-1474.

2 Zgurskaya, H. I.; Krishnamoorthy, G.; Tikhonova, E. B.; Lau, S. Y.; Stratton, K. L. Mechanism of antibiotic efflux in Gram-negative bacteria. *Frontiers in Bioscience* **2003**, *8*, S862-S873.

3 Nies, D. H. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol. Rev.* **2003**, *27*, 313-339.

4 Hvorup, R. N.; Winnen, B.; Chang, A. B.; Jiang, Y.; Zhou, X.; Saier, M. H. The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. *Eur. J. Biochem.* **2003**, *270*, 799-813.

5 (a) Bost, S.; Silva, F.; Belin, D. Transcriptional activation of *ydeA*, which encodes a member of the major facilitator superfamily, interferes with arabinose accumulation and induction of the *Escherichia coli* arabinose $P_{BAD}$ promoter. *J. Bacteriol.* **1999**, *181*, 2185-2191. (b) Liu, J. Y.; Miller, P. F.; Willard, J.; Olson, E. R. Functional and biochemical characterization of *Escherichia coli* sugar efflux transporters. *J. Biol. Chem.* **1999**, *274*, 22977-22984.

6 (a) Aleshin, V. V.; Zakataeva, N. P.; Livshits, V. A. A new family of amino-acid-efflux proteins. *Trends Biochem. Sci.* **1999**, *24*, 133-135. (b) Dabler, T.; Maier, W; Winterhalter, C.; Bock, A. Identification of a major facilitator protein from *Escherichia coli* involved in efflux of metabolites of the cysteine pathway. *Mol. Microbiol.* **2000**, *36*, 1101-1112. (c)Vrljic, M.; Sahm, H.; Eggeling, L. The LysE superfamily: topology of the lysine exporter LysE of *Corynebacterium glutamicum*, a paradigm for a novel superfamily of transmembrane solute translocators. *J. Mol. Microbiol. Biotechnol.* **1999**, *1*, 327-336. (d) Eggeling, L.; Sahm, H. New ubiquitous translocators: amino acid export by *Corynebacterium glutamicum* and *Escherichia coli*. *Arch. Microbiol.* **2003**, *180*, 155-160.

7 Bellmann, A. M.; Vrljic, M.; Patek, M.; Sahm, H.; Kramer, R.; Eggeling, L. Regulation and specificity of the lysE-mediated export of amino acids by *Corynebacterium glutamicum*. *Microbiology* **2001**, *147*, 1765-1774.

8 Vrljic, M.; Sahm, H.; Eggeling, L. A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. *Mol. Microbiol.* **1996**, *22*, 815-826.

9 Nisbet, T. M.; Payne, J. W. The characteristics of peptide uptake in *Streptococcus faecalis*: studies on the transport of natural peptides and antibacterial phosphonopeptides. *J. Gen. Microbiol.* **1982**, *128*, 1357-1364.

10 Payne, J. W.; Bell, G. Direct determination of the properties of peptide transport systems in *Escherichia coli*, using a fluorescent-labeling procedure. *J. Bacteriol.* **1979**, *137*, 447-455.

11 Juillard, V.; Le Bars, D.; Kunji, E. R. S.; Konings, W. N.; Gripon, J.; Richard, J. Oligopeptides are the main source of nitrogen for *Lactococcus lactis* during growth on milk. *Appl. Environ. Microbiol.* **1995**, *61*, 3024-3030.

12 Kunji, E. R. S.; Mierau, I.; Poolman, B.; Konings, W. N.; Venema, G.; Kok, J. Fate of peptides in peptidase mutants of *Lactococc us lactis. Mol. Microbiol.* **1996**, *21*, 123-131.

13 Kramer, R. Secretion of amino acid by bacteria: physiology and mechanism. *FEMS Microbiol. Rev.* **1994**, *13*, 75-94.

14 Simic, P.; Sahm, H.; Eggeling, L. L-Threonine Export: Use of Peptides To Identify a New Translocator from *Corynebacterium glutamicum. J. Bacteriol.* **2001**, *183*, 5317-5324.

15 Zakataeva, N. P.; Aleshin, V. V.; Tokmakova, I. L.; Troshin, P. V.; Livshits, V. A. The Novel Transmembrane *Escherichia coli* Proteins Involved in the Amino Acid Efflux. *FEBS Lett.* **1999**, *452*, 228-232.

16 Vrljic, M.; Sahm, H.; Eggeling, L. A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum. Mol. Microbiol.* **1996**, *22*, 815-826.

17 Morbach, S.; Sahm, H.; Eggeling, L. L-Isoleucine Production with *Corynebacterium glutamicum*: Further Flux Increase and Limitation of Export. *Appl. Environ. Microbiol.* **1996**, *62*, 4345-4351.

18 (a) Kruse, D.; Kramer, R.; Eggeling, L.; Rieping, M.; Pfefferle, W.; Tchieu, J. H.; Chung, Y. J.; Jr Saier, M. H.; Burkovski, A. Influence of threonine exporters on threonine production in *Escherichia coli. Appl. Microbiol. Biotechnol.* **2002**, *59*, 205-210. (b) Simic, P.; Willuhn, J.; Sahm, H.; Eggeling, L. Identification of *glyA* (Encoding Serine Hydroxymethyltransferase) and Its Use Together with the Exporter ThrE To Increase L-Threonine Accumulation by *Corynebacterium glutamicum. Appl. Environ. Microbiol.* **2002**, *68*, 3321-3327.

19 (a) Rottenberg, H. The measurement of membrane potential and pH in cells, organelles, and vesicles. *Methods Enzymol.* **1979**, *55*, 547-569. (b) Kashket, E. R. The proton motive force in bacteria: A critical assessment of methods. *Annu. Rev. Microbiol.* **1985**, *39*, 219-242. (c) Ebbighausen, H.; Weil, B.; Kramer, R. Isoleucine excretion in *Corynebacterium glutamcum*: evidence for a specific efflux carrier system. *Appl. Microbiol. Biotechnol.* **1989**, *31*, 184-190. (d) Zakataeva, N.; Aleshin, V. V.; Tokmakova, I. L.; Troshin, P. V.; Livshits, V. A. *FEBS Lett.* **1999**, *452*, 228-232.

20 Cromartie, T. H.; Polge, N. D. In *U.S.*; (Syngenta Limited, UK). Method of Detecting Shikimic Acid. Int. Patent WO00/49399, August 24, 2000.

21 Chandran, S. S.; Yi, J.; Draths, K. M.; von Daeniken, R.; Weber, W.; Frost, J. W. Phosphoenolpyruvic Acid Availability and the Biosynthesis of Shikimic Acid. *Biotechol. Prog.* **2003**, *19*, 808-814.

22 (a) Klingenberg, M.; Pfaff, E. Means of terminating reactions. *Methods Enzymol.* **1967**, *10*, 680-684. (b) Ebbighausen, H.; Weil, B.; Kramer, R. Transport of branched-chain amino acids in *Corynebacterium glutamicum. Arch. Microbiol.* **1989**, *151*, 238-244.

23 Draths, K. M. Unpublished results.

24 Kleckner, N.; Bender, J.; Gottesman, S. Uses of Transposons with Emphasis on Tn10. *Meth. Enzymol.* **1991**, *204*, 139-180.

25 Miller, J. H. *A short course in bacterial genetics*; Cold Spring Harbor Laboratory: Plainview, NY, 1992.

26 Kwon, Y. M.; Ricke, S. C. Efficient Amplification of Multiple Transposon-Flanking Sequences. *J. Microbiol. Methods* **2000**, *41*, 195-199.

27 Clegg, S.; Yu, F.; Griffiths, L.; Cole, J. A. The roles of the polytopic membrane proteins NarK, NarU and NirC in *Escherichia coli* K-12: two nitrate and three nitrite transporters. *Mol. Microbiol.* **2002**, *44*, 143-155.

28 Yi, J.; Li, K.; Draths, K. M.; Frost, J. W. Modulation of Phosphoenolpyruvate Synthase Expression Increases Shikimate Pathway Product Yields in *E. coli*. *Biotechnol. Prog.* **2002**, *18*, 1141-1148.

29 Davies, G. M.; Barrett-Bee, K. J.; Jude, D. A.; Lehan, M.; Nichols, W. W.; Pinder, P. E.; Thain, J. L.; Watkins, W. J.; Wilson, R. G. (*6S*)-6-Fluoroshikimic Acid, an Antibacterial Agent Acting on the Aromatic Biosynthetic Pathway. *Antimicrob. Agents and Chemo.* **1994**, *38,* 403-406.

30 Ewart, C. D. C.; Jude, D. A.; Thain, J. L.; Nichols, W. W. Frequency and Mechanism of Resistance to Antibacterial Action of ZM240401, (6*S*)-6-Fluoro-Shikimic Acid. *Antimicrob. Agents and Chemo.* **1995**, *39,* 87-93.

31 Tian, F. Mechanism-Based Inhibition of 3-Dehydroquinate Synthase and *myo*-Inositol 1-Phosphate Synthase. Ph. D. Thesis, Michgan State University, 1998.

32 (a) Sutherland, J. K. Watkins, W. J. Bailey, J. P.; Chapman, A. K. Davies, G. M. The Synthesis of 6α- and 6β-Fluoroshikimic Acids. *J. Chem. Soc. Chem. Commun.* **1989**, *18*, 1386-1387. (b) Bowles, S.; Campbell, M. M.; Sainsbury, M.; Davies, G. M. Reactivity Studies in the Shikimic Acid Series. *Tetrahedron Lett.* **1989**, *30*, 3711-3714. (c) Bowles, S. A.; Campbell, M. M.; Sainsbury, M.; Davies, G. M. Reactivity Studies in the Shikimic Acid Series: The Synthesis of Racemic Methyl 6α-Fluoroshikimate. *Tetrahedron* **1990**, *46*, 3981-3992.

33 Pitcher, D. G.; Saunders, N. A.; Owen, R. J. Rapid Extraction of Bacterial Genomic DNA with Guanidium Thiocyanate. *Letters in Applied Microbiology* **1989**, *8*, 151-156.

34 Bornemann, S.; Ramjee, M. K.; Balasubramnian, S.; Abell, C.; Coggins, J. R.; Lowes, D. J.; Thorneley, R. N. F. *Escherichia coli* Chorismate Synthase Catalyzes the Conversion of (6*S*)-6-Fluoro-5-enolpyruvylshikimate-3-phosphate to 6-Fluorochorismate. *J. Biol. Chem.* **1995**, *270*, 22811-22815.

35 Bornemann, S.; Ramjee, M. K.; Lowe, D. J.; Thorneley, R. N. F.; Coggins, J. R.; Abell, C.; Balasubramanian, S.; Hawkes, T. R.; Nichols, W. W.; Davies, G. M. Studies on *Escherichia coli* chorismate synthase. p. 843-846. *In* K. Yagi (ed.), Flavins and flavoproteins 1993. Walter de Gruyter, Berlin.

# CHAPTER 5

# Experimental

## General methods

### Spectroscopic measurements

$^1$H NMR spectra were recorded on a Varian 300 MHz VX-300 FT-NMR spectrometer. Chemical shifts were reported in parts per million (ppm) downfield from internal sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$ (TSP, $\delta$ = 0.00) with $D_2O$ as solvent. TSP was purchased from Lancaster. UV and visible measurements were recorded on a Perkin-Elmer Lambda 3b UV-vis spectrophotometer or on a Hewlett Packard 8452A Diode Array Spectrophotometer equipped with HP 89532A UV-Visible Operating Software.

### Bacteria strains and plasmids

All the strains and plasmids used are shown in Table 31. *E. coli* K-12 strain RB791 was obtained from the American Type Culture Collection (ATCC strain 53622). *E. coli* LE392 was obtained from the American Type Culture Collection (ATCC number 53010). *E. coli* AB2834,[1] JM2053, AB2847[1] were obtained from the *E. coli* Genetic Stock Center at Yale University. *E. coli* KL3),[2] SP1.1,[3] SP3.1 and SP1.4 were constructed in the lab previously. *E. coli* AL0807 was provided by Professor M. G. Marinus (University of Massachusetts). *E. coli* UE79[4] was provided by Professor W. Boos (University of Konstanz). *E. coli* CB734[5] was provided by Professor Akowski.

**Table 31. Bacterial strains and plasmids.**

| Strain/Plasmid | Relevant Characteristics | Source |
|---|---|---|
| | Strain | |
| RB791 | W3110 *lacL8I$^q$* | ATCC |
| LE392 | *F$^-$ supE44 supF58 lacY1 galK2 galT22 metB1 trpR55 hsdR514 mcrB λ* | ATCC |
| AB2834 | *tsx-352 glnV42 λ$^-$ aroE353 malT352* | CGSC |
| JM2053 | *ptsG0 manY0(Am) uhp-0(Const) fruA0 thyA0 galP77::Tn10 ilvD kga0 nag0 mglP0 umgC0 pel0(Am)* | CGSC |
| AB2847 | *tsx-354 glnV42 λ$^-$ aroB351 malT354* | CGSC |
| KL3 | AB2834 *serA::aroB* | Ref. 2 |
| SP1.1 | RB791 *serA::aroB aroL478::Tn10 aroK17::Cm$^R$* | Ref. 3 |
| SP3.1 | KL3 *aroL478::Tn10 aroK17::Cm$^R$* | Lab |
| SP1.4 | RB791 *serA::aroF$^{FBR}$aroB aroL::Tn10 tyrR3* | Lab |
| AL0807 | *F$^-$ aroL478::Tn10 aroK17::Cm$^R$* | Marinus |
| UE79 | *galR glK thi ΔptsHIcrr zfc-706::Tn10* | Ref. 4 |
| CB734 | C600 *leu thi1 Δ(gal-aroG-nadA)50 aroF:cat (Cm$^r$) ΔaroH::Kan$^r$ recA1* | Ref. 5 |
| XL-10 Gold | *{Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacI$^q$ZΔM15 Tn10 Amy Cam$^r$]* | Stratagene |
| DH5α | *F$^-$ ϕ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17($r_{k-}$, $m_{k+}$) phoA λ$^-$ supE44 thi-1 gyrA96 relA1* | Invitrogen |
| | Plasmid | |
| pJF118EH | Ap$^R$, $P_{tac}$ *lacI$^q$* | Ref. 6 |
| pTC325 | $P_{tac}$*glfglk* | Ref. 7 |
| pD2625 | *serA* source | GCI |
| pSU18 | Cm$^R$, $P_{lac}$*lacZ'*, p 15A replicon | Ref. 8 |
| pBR322 | Ap$^R$, Tet$^R$, *rep* replicon | Ref. 9 |
| p34H | Ap$^R$ | Ref. 10 |
| pKL1.87A | Cm$^R$, *ppsA* | Lab |
| pKL4.66B | Cm$^R$, *aroF$^{FBR}$*, *aroF$^{FBR}$*, *serA* in pSU18 | Ref. 2 |
| pKD11.291A | Cm$^R$, *serA*, *aroF$^{FBR}$*, $P_{aroF}$ | Ref. 2 |
| pKL5.17A | Cm$^R$, *tktA* in pKD11.291A | Ref. 2 |
| pMF51A | Ap$^R$, *tktA* | Lab |
| pMF63A | Ap$^R$, *aroF$^{FBR}$* | Lab |
| pKD12.138 | Ap$^R$, *aroF$^{FBR}$*, *tktA*, $P_{tac}$*aroE*, *serA* | Ref. 11 |
| pKD15.071B | *ppsA*, *aroF$^{FBR}$*, *tktA*, $P_{tac}$*aroE*, *serA* | Ref. 11 |

*E. coli* XL10-Gold was obtained from Stratagene. Plasmid constructions were carried out in *E. coli* DH5, which is available from Invitrogen.

Plasmid pJF118EH[6] was provided by Professor M. Bagdasarian (Michigan State University). Plasmid pTC325[7] was provided by Professor Lonnie O. Ingram (University of Florida). Plasmid pD2625 was obtained from Genencor International. Plasmids pSU18,[8] pBR322,[9] and p34H[10] were obtained previously by this lab. Plasmids pKL1.87A, pKL4.66B,[2] pKD11.291A,[2] pKL5.17A,[2] pMF51A,[11] pMF63A,[11] pKD12.138,[12] pKD15.071B[12] were constructed in the lab previously.

**Storage of bacterial strains and plasmids**

All bacterial strains were stored at -78 °C in glycerol. Plasmids were transformed into DH5α for long-term storage. Glycerol samples were prepared by adding 0.75 mL of an overnight culture to a sterile vial containing 0.25 mL of 80% (v/v) glycerol. The solution was mixed, left at room temperature for 2 h and then stored at -78 °C.

**Culture medium**

All solutions were prepared in distilled, deionized water. LB medium[13] (1 L) contained Bacto tryptone (10 g), Bacto yeast extract (5 g), and NaCl (10 g). L-Broth[13] (1 L) contained Bacto tryptone (10 g), Bacto yeast extract (5 g), NaCl (5 g), glucose (1 g) and $CaCl_2$ (2.5 mM). Soft agar[13] (100 mL) contained Bacto tryptone (1 g), Difco agar (0.55 g), and NaCl (0.5 g). TB medium[13] (1 L) contained tryptone (10 g) and NaCl (5 g). After autoclaving and directly before use, $MgSO_4$ (10 mL of 1 M stock per L) was added to the TB medium. M9 salts[13] (1 L) contained $Na_2HPO_4$ (6 g), $KH_2PO_4$ (3 g), $NH_4Cl$ (1 g), and NaCl (0.5 g). M9 medium contained carbon sources (D-glucose, D-xylose, D-maltose or D-mannitol, 10 g), $MgSO_4$ (0.12 g), and thiamine (0.001 g) in 1 L of M9 salts. Solutions of inorganic salts, magnesium salts, and carbon sources were autoclaved separately and then mixed. Antibiotics were added where appropriate to the following final concentrations unless noted otherwise: chloramphenicol, 20 μg/mL; ampicillin, 50 μg/mL; kanamycin, 50 μg/mL; tetracycline, 12.5 μg/mL; and fosfomycin, 20 μg/mL. Stock solution of antibiotics were prepared in water with the exception of chloramphenicol which was prepared in 95% ethanol and tetracycline which was prepared in 50% aqueous ethanol. L-Phenylalanine, L-tyrosine, L-tryptophan, and L-

serine were added to M9 medium where indicated to a final concentration of 0.04 g/L. Antibiotics, isopropyl β-D-thioglucopyranoside (IPTG), thiamine, cAMP and amino acid supplementations were sterilized through 0.22-μm membranes prior to addition to M9 medium. Solid medium was prepared by addition of 1.5% (w/v) Difco agar to the medium. Fermentation medium (1 L) contained $K_2HPO_4$ (7.5 g), ammonium iron (III) citrate (0.3 g), citric acid monohydrate (2.1 g), L-phenylalanine (0.7 g), L-tyrosine (0.7 g), and L-tryptophan (0.35 g), and concentrated $H_2SO_4$ (1.2 mL). The culture medium was adjusted to pH 7.0 by addition of concentrated $NH_4OH$ before autoclaving. The following supplementations were added immediately prior to initiation of the fermentation: glucose (19-24 g under glucose-limited conditions or 30 g under glucose-rich conditions), $MgSO_4$ (0.24 g), aromatic vitamins *p*-aminobenzoic acid (0.01 g), 2,3-dihydroxybenzoic acid (0.01 g), and *p*-hydroxybenzoic acid (0.01 g), and trace minerals $(NH_4)_6(Mo_7O_{24})\cdot 4H_2O$ (0.0037 g), $ZnSO_4\cdot 7H_2O$ (0.0029 g), $H_3BO_3$ (0.0247 g), $CuSO_4\cdot 5H_2O$ (0.0025 g), and $MnCl_2\cdot 4H_2O$ (0.0158 g). D-Glucose and $MgSO_4$ were autoclaved separately while aromatic vitamins and trace minerals were sterilized through 0.22-μm membranes prior to addition to the medium.

**Fed-batch fermentation (general)**

Fermentations[14] employed a 2.0 L working capacity B. Braun M2 culture vessel fitted with a stainless steel baffle cage consisting of four 1/2" x 5" baffles. Utilities were supplied by a B. Braun Biostat MD controlled by a DCU-1 or DCU-3. Data acquisition utilized a Dell Optiplex $Gs^+$ 5166M personal computer (PC) equipped with B. Braun MFCS/Win software (v2.0). Temperature, pH, and dissolved oxygen (D.O) were

controlled with PID control loops. Temperature was maintained at 36 °C, and pH was maintained at 7.0 by addition of concentrated $NH_4OH$ or 2 *N* $H_2SO_4$. Dissolved oxygen was measured using a Mettler-Toledo 12 mm sterilizable $O_2$ sensor fitted with an Ingold A-type $O_2$ permeable membrane. D.O. was maintained at 20% air saturation. Antifoam (Sigma 204) was added as needed.

Inoculants were prepared by introduction of a single colony into 5 mL of M9 medium. The culture was grown at 37 °C with agitation at 250 rpm until they were turbid (~18-30 h) and subsequently transferred to 100 mL of M9 medium. Cultures were grown at 37 °C for an additional 12 h. The inoculant ($OD_{600}$ = 1.0-2.0) was then transferred into the fermentor vessel and the batch fermentation was initiated (t = 0 h).

**Glucose-rich fermentor conditions**

The initial glucose concentration in the fermentation medium was 30 g/L. Three staged methods were used to maintain D.O. levels at 20% air saturation during the course of the fermentations. With the airflow at an initial setting of 0.06 L/L/min, D.O. concentration was maintained by increasing the impeller speed from its initial set point of 50 rpm to a preset maximum of 750 rpm. With the impeller rate constant at 750 rpm, the mass flow controller then maintained D.O. levels by increasing the airflow rate from 0.06 L/L/min to a preset maximum of 1.0 L/L/min. After the preset maxima of 750 rpm and 1.0 L/L/min were reached, the third stage of the fermentation was initiated in which glucose (65% w/v) was added to the vessel at a rate sufficient to maintain a glucose concentration in the range of 5 to 30 g/L for the remainder of the run. Airflow was maintained at 1.0 L/L/min, and the impeller was allowed to vary in order to maintain the

D.O. concentration at 20% air saturation. The impeller speed typically varied from 750 rpm to 1400 rpm during the remainder of the run. A solution of IPTG (100 mM; 0, 0.25, 0.50, 0.75, 1.0, and 2.0 mL) was added at timed intervals after initiation of the run to achieve reported IPTG concentrations respectively of 0, 6.0, 12, 18, 24, and 48 mg/L in the fermentation medium.

**Glucose-limited fermentor conditions**

The initial glucose concentration in the fermentation medium was 19-24 g/L, depending on the strain being examined. Three staged methods were used to maintain D.O. levels at 20% air saturation, with the first two stages identical to those described for the glucose-rich conditions. After the preset maxima of 750 rpm and 1.0 L/L/min of airflow were reached, the third stage of the fermentation was initiated in which the D.O. concentration was maintained at 20% air saturation for the remainder of the run by oxygen sensor-controlled glucose feeding. At the beginning of this stage, the D.O. concentration initially fell below 20% air saturation due to residual initial glucose in the medium. This lasted for up to 30 min before glucose (65% w/v) feeding commenced. The glucose feed PID control parameters were set to 0.0 s (off) for the derivative control ($\tau_D$) and 999.9 s (minimum control action) for the integral control ($\tau_I$). $X_P$ was set to 950% to achieve a $K_c$ of 0.1. A solution of IPTG (100 mM; 0, 0.25, 0.50, 0.75, 1.0, and 2.0 mL) was added at timed intervals after initiation of the run to achieve reported IPTG concentrations respectively of 0, 6.0, 12, 18, 24, and 48 mg/L in the fermentation medium.

**Analysis of fermentation broths**

Samples (5 mL) of fermentation broth were taken at the indicated timed intervals. Cell densities were determined by dilution of fermentation broth with water (1:100) followed by measurement of absorption at 600 nm ($OD_{600}$). Dry cell weight (g/L) was calculated using a conversion coefficient of 0.43 g/L/$OD_{600}$. The remaining fermentation broth was centrifuged to obtain cell-free broth.

Glucose concentrations in cell-free broth were measured using the Glucose Diagnostic Kit purchased from Sigma. Solute concentrations in the cell-free broth were quantified by $^1$H NMR. Solutions were concentrated to dryness under reduced pressure, concentrated to dryness one additional time from $D_2O$, and then redissolved in $D_2O$ containing a known concentration of the sodium salt of 3-(trimethylsilyl)propionic-2,2,3,3-$d_4$ acid (TSP). $^1$H NMR spectra were recorded and concentrations were determined by comparison of integrals corresponding to each compound with the integral corresponding to TSP ($\delta$ = 0.00 ppm). A standard concentration curve was determined for each metabolite using solutions of authentic, purified metabolites. The following resonances were used to quantify each compound: shikimic acid ($\delta$ 4.57, d, 1 H); 3-dehydroshikimic acid ($\delta$ 4.28, d, 1 H); 3-dehydroquinic acid ($\delta$ 4.38, d, 1 H); 3-deoxy-D-*arabino*-heptulosonic acid ($\delta$ 1.81, t, 1 H); and gallic acid ($\delta$ 7.02, s, 2 H). The following response factor was used for each molecule: shikimic acid, 0.70; 3-dehydroshikimic acid, 0.95; 3-dehydroquinic acid, 0.89; 3-deoxy-D-*arabino*-heptulosonic acid, 1.22; gallic acid, 1.36.

## Genetic manipulations

### General

Recombinant DNA manipulations generally followed methods described by Sambrook.[15] Restriction enzymes were purchased from Invitrogen or New England Biolabs. T4 DNA ligase was obtained from Invitrogen. Fast-Link™ DNA Ligation Kit was obtained from Epicentre. Zymoclean Gel DNA Recovery Kit and DNA Clean & Concentrator Kit was obtained from Zymo Research Company. Maxi and Midi Plasmid Purification Kits were obtained from Qiagen. Calf intestinal alkaline phosphatase was obtained from Boehringer Mannheim. Agarose (electrophoresis grade) was obtained from Invitrogen. Phenol was prepared by addition of 0.1 % (w/v) 8-hydroxyquinoline to distilled, liquefied phenol. Extraction with an equal volume of 1 M Tris-HCl (pH 8.0) two times was followed by extraction with 0.1 M Tris-HCl (pH 8.0) until the pH of the aqueous layer was greater than 7.6. Phenol was stored at 4 °C under an equal volume of 0.1 M Tris-HCl (pH 8.0). SEVAG was a mixture of chloroform and isoamyl alcohol (24:1 v/v). TE buffer contained 10 mM Tris-HCl (pH 8.0) and 1 mM $Na_2EDTA$ (pH 8.0). Endostop solution (10X concentration) contained 50% glycerol (v/v), 0.1 M $Na_2EDTA$, pH 7.5, 1% sodium dodecyl sulfate (SDS) (w/v), 0.1% bromophenol blue (w/v), and 0.1% xylene cyanole FF (w/v) and was stored at 4 °C. Prior to use, 0.12 mL of DNase-free RNase was added to 1 mL of 10X Endostop solution. DNase-free RNase (10 mg $mL^{-1}$) was prepared by dissolving RNase in 10 mM Tris-Cl (pH 7.5) and 15 mM NaCl. DNase activity was inactivated by heating the solution at 100 °C for 15 min. Aliquots were stored at -20 °C. PCR amplifications were carried out as described by Sambrook.[15] Each reaction (0.1 mL) contained 10 mM KCl, 20 mM Tris-Cl (pH 8.8), 10

mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, dATP (0.2 mM), dCTP (0.2 mM), dGTP (0.2 mM), dTTP (0.2 mM), template DNA, 0.5 μM of each primer, and 2 units of Vent. Taq or Pfu polymerase also have been used for PCR reaction with the reaction buffers provided. Initial template concentrations varied from 0.02 μg to 1.0 μg.

**Large scale purification of plasmid DNA**

Plasmid DNA was purified on a large scale using a modified alkaline lysis method described by Sambrook.[15] In a 2 L Erlenmeyer flask, LB (500 mL) containing the appropriate antibiotics was inoculated from a single colony, and the culture was incubated in a gyratory shaker (250 rpm) for 14 h at 37 °C. Cells were harvested by centrifugation (4 000*g*, 5 min, 4 °C) and then resuspended in 10 mL of cold GETL solution [50 mM glucose, 20 mM Tris-HCl (pH 8.0), 10 mM $Na_2EDTA$ (pH 8.0)] into which lysozyme (5 mg $mL^{-1}$) had been added immediately before use. The suspension was stored at room temperature for 5 min. Addition of 20 mL of 1% sodium dodecyl sulfate (w/v) in 0.2 N NaOH was followed by gentle mixing and storage on ice for 15 min. Fifteen milliliters of an ice cold solution containing 3 M KOAc (prepared by combining 60 mL of 5 M potassium acetate, 11.5 mL of glacial acetic acid, and 28.5 mL of $H_2O$) was added. Vigorous shaking resulted in formation of a white precipitate. After the suspension was stored on ice for 10 min, the cellular debris was removed by centrifugation (48 000*g*, 20 min, 4 °C). The supernatant was transferred to two clean centrifuge bottles and isopropanol (0.6 volumes) was added to precipitate the DNA. After the samples were left at room temperature for 15 min, the DNA was recovered by

centrifugation (20 000*g*, 20 min, 4 °C). The DNA pellet was then rinsed with 70% ethanol and dried.

Further purification of the DNA sample involved precipitation with polyethylene glycol (PEG). The DNA was dissolved in TE (3 mL) and transferred to a Corex tube. Cold 5 M LiCl (3 mL) was added and the solution was gently mixed. The sample was then centrifuged (12 000*g*, 10 min, 4 °C) to remove high molecular weight RNA. The clear supernatant was transferred to a clean Corex tube and isopropanol (6 mL) was added followed by gentle mixing. The precipitated DNA was collected by centrifugation (12 000*g*, 10 min, 4 °C). The DNA was then rinsed with 70% ethanol and dried. After redissolving the DNA in 0.5 mL of TE containing 20 μg/mL of RNase, the solution was transferred to a 1.5 mL microcentrifuge tube and stored at room temperature for 30 min. To this sample was added 500 μL of 1.6 M NaCl containing 13% PEG-8000 (w/v) (Sigma). The solution was mixed and centrifuged (microcentrifuge, 10 min, 4 °C) to recover the precipitated DNA. The supernatant was removed and the DNA was then redissolved in 400 μL of TE. The sample was extracted sequentially with phenol (400 μL), phenol and SEVAG (400 μL each), and finally SEVAG (400 μL). Ammonium acetate (10 M, 100 μL) was added to the aqueous DNA solution. After thorough mixing, 95% ethanol (1 mL) was added to precipitate the DNA. The sample was left at room temperature for 5 min and then centrifuged (microcentrifuge, 5 min, 4 °C). The DNA was rinsed with 70% ethanol, dried, and then redissolved in 200-500 μL of TE.

Alternatively, DNA was purified using a Qiagen Maxi Kit or Midi Kit as described by the manufacturer. The purity of DNA isolated by this method was adequate for DNA sequencing.

**Small scale purification of plasmid DNA**

An overnight culture (5 mL) of the plasmid-containing strain was grown in LB containing the appropriate antibiotics.[15] Cells from 3 mL of the culture were collected in a 1.5 mL microcentrifuge tube by centrifugation. The resulting cell pellet was liquefied by vortexing (30 sec) and then resuspended in 0.1 mL of cold GETL solution into which lysozyme (5 mg $mL^{-1}$) had been added immediately before use. The solution was stored on ice for 10 min. Addition of 0.2 mL of 1% sodium dodecyl sulfate (w/v) in 0.2 N NaOH was followed by gentle mixing and storage on ice for 5-10 min. To the sample was added 0.15 mL of cold KOAc solution. The solution was shaken vigorously and stored on ice for 5 min before centrifugation (15 min, 4 °C). The supernatant was transferred to another microcentrifuge tube and extracted with equal volumes of phenol and SEVAG (0.2 mL). The aqueous phase (approximately 0.5 mL) was transferred to a fresh microfuge tube, and DNA was precipitated by the addition of 95% ethanol (1 mL). The sample was left at room temperature for 5 min before centrifugation (15 min, room temperature) to collect the DNA. The DNA pellet was rinsed with 70% ethanol, dried, and redissolved in 50 -100 μL TE. DNA isolated from this method was used for restriction enzyme analysis, and the concentration was not determined by spectroscopic methods.

**Determination of DNA concentration**

The concentration of DNA in the sample was determined as follows. An aliquot (10 μL) of the DNA was diluted to 1 mL in TE and the absorbance at 260 nm was

measured relative to the absorbance of TE. The DNA concentration was calculated based on the fact that the absorbance at 260 nm of a 50 μg $mL^{-1}$ of plasmid DNA is 1.0.

**DNA precipitation**

DNA was precipitated by addition of 0.1 volume of 3 M NaOAc (pH 5.2) followed by thorough mixing and addition of 3 volumes of 95% ethanol. Samples were stored for at least 2 h at -78 °C. Precipitated DNA was recovered by centrifugation (15 min, 4 °C). To the DNA pellet was added 70% ethanol (100 μL), and the sample was centrifuged again (15 min, 4 °C). DNA was dried and redissolved in TE.

**Restriction enzyme digestion of DNA**

Restriction enzyme digests were performed using restriction enzyme buffers supplied by Invitrogen or New England Biolabs. A typical digest contained approximately 0.8 μg of DNA in 8 μL TE, 2 μL of restriction enzyme buffer (10X concentration), 1 μL of restriction enzyme, and TE to a final volume of 20 μL. Reactions were incubated at 37 °C for 1 h. Digests were terminated by addition of 2.2 μL of Endostop solution (10X concentration) and subsequently analyzed by agarose gel electrophoresis. When DNA was required for subsequent cloning, restriction digests were terminated by addition of 1 μL of 0.5 M $Na_2EDTA$ (pH 8.0) followed by extraction of the DNA with equal volumes of phenol and SEVAG and precipitation of the DNA.

**Agarose gel electrophoresis**

Agarose gels were run in TAE buffer containing 40 mM Tris-acetate and 2 mM EDTA (pH 8.0). Gels typically contained 0.7% agarose (w/v) in TAE buffer. Higher concentrations of agarose (1%-2%) were used to resolve DNA fragments smaller than 1 kb. Lower concentrations of agarose (0.35%) were used to resolve DNA fragments larger than 10 kb. Ethidium bromide (0.5 $\mu g\ mL^{-1}$) was added to the agarose to allow visualization of DNA fragments over a UV lamp. The size of the DNA fragments were determined by using two sets of DNA standards: λ DNA digested with *Hin*dIII (23.1-kb, 9.4-kb, 6.6-kb, 4.4-kb, 2.3-kb, 2.0-kb, and 0.6-kb) and λ DNA digested with *Eco*RI and *Hin*dIII (21.2-kb, 5.1-kb, 5.0-kb, 4.3-kb, 3.5-kb, 2.0-kb, 1.9-kb, 1.6-kb, 1.4-kb, 0.9-kb, 0.8-kb, and 0.6-kb). Also 100 bp DNA Ladder (Invitrogen) was used to determine the size of small DNA fragements. The ladder consists of 15 blunt-ended fragments ranging in length from 100 to 1500 bp, at 100 bp increments, and an additional fragment at 2,072 bp.

**Isolation of DNA from agarose**

The band of agarose containing DNA of interest was excised from the gel while visualized with high wavelength UV and chopped thoroughly with a razor in a plastic weighing tray. The agarose was then transferred to a spin column consisting of a 500 μL microfuge tube packed tightly with glass wool and with an 18 gauge hole in its bottom. The spin column was then centrifuged for 5 min using a microcentrifuge to separate the DNA solution from the agarose. The DNA-containing aqueous phase collected after

centrifugation were mixed with 3 M NaOAc and 95% ethanol. The DNA was precipitated as described previously and redissolved in TE.

Alternatively, the band of agarose containing DNA of interest was excised from the gel while visualized with high wavelength UV. Zymoclean Gel DNA Recovery Kit was used to isolate DNA from the agarose gel according to the protocol provided by the Zymo Research Company.

**Treatment of vector DNA with calf intestinal alkaline phosphatase**

Plasmid vectors digested with a single restriction enzyme were dephosphorylated to prevent self-ligation. Vector DNA after digestion was immediately combined in a total volume of 60 μL. To this sample was added 7 μL of dephosphorylation buffer (10X concentration, provided by enzyme supplier) and 3 μL of calf intestinal alkaline phosphatase (3 units). The reaction was incubated at 37 °C for 1 h. The phosphatase was inactivated by addition of 1 μL of 0.5 M $Na_2EDTA$ (pH 8.0) followed by heat treatment (65 °C, 20 min). The sample was extracted with phenol and SEVAG (100 μL each) to remove protein, and the DNA was precipitated as previously described and redissolved in TE.

**Treatment of DNA with Klenow fragment**

DNA with recessed 3' termini was modified to blunt-ended fragment by treatment with the Klenow fragment of *E. coli* DNA polymerase I. After the DNA (0.8-2 μg) restriction digestion was completed in a 20 μL reaction, a solution (1 μL) containing each of the desired dNTPs was added to provide a final concentration of 1 mM for each dNTP.

Addition of 1-2 units of the Klenow fragment to the reaction was followed by incubation of the mixture at room temperature for 20-30 min. Since the Klenow fragment works well in the common buffers used for restriction digestion of DNA, there was no need to purify the DNA after restriction digestion and prior to filling recessed 3' termini. Klenow reactions were quenched by extraction with equal volumes of phenol and SEVAG. DNA was recovered by DNA precipitation.

**Ligation of DNA**

DNA ligations were designed so that the molar ratio of insert to vector was 3 to 1. A typical ligation reaction contained 0.03 to 0.1 μg of vector and 0.05 to 0.2 μg of insert in a total volume of 7 μL. To this sample was added 2 μL of ligation buffer (5X concentration, provided by the enzyme supplier) and 1 μL of T4 DNA ligase (2 units). The reaction was incubated at 16 °C for at least 4 h and then used to transform competent cells.

Alternatively, Fast-Link DNA Ligation Kit (Epicentre) was used for ligation of insert DNA with cohesive or blunt ends into vectors with compatible cohevsive ends according to the protocol provided by the manufacturer.

**Preparation and transformation of competent cells**

Competent cells were prepared using a procedure modified from Sambrook.[15] An aliquot (1 mL) from an overnight culture (5 mL) was used to inoculate 100 mL of LB (500 mL Erlenmeyer flask) containing the appropriate antibiotics. The cells were cultured in a gyratory shaker (37 °C, 250 rpm) until they reached the mid-log phase of

growth (judged from the absorbance at 600 nm reaching 0.4-0.6). The culture was poured into a large centrifuge bottle that had been previously sterilized with bleach and rinsed with sterile water. The cells were collected by centrifugation (4 000*g*, 5 min, 4 °C) and the culture medium was discarded. All manipulations were carried out on ice during the remaining portion of the procedure. The cell pellet was washed with 100 mL of cold 0.9% NaCl (w/v) and then resuspended in 50 mL of cold 100 mM $CaCl_2$. The suspension was stored on ice for a minimum of 30 min and then centrifuged (4 000*g*, 5 min, 4 °C). The cell pellet was resuspended in 4 mL of cold 100 mM $CaCl_2$ containing 15% glycerol (v/v). Aliquots (0.25 mL) were dispensed into 1.5 mL microcentrifuge tubes and immediately frozen in liquid nitrogen. Competent cells were stored at -78 °C with no significant decrease in transformation efficiency over a period of six months.

Frozen competent cells were thawed on ice for 5 min before transformation. A small aliquot (1 to 10 μL) of plasmid DNA or a ligation reaction was added to the thawed competent cells (0.1 mL). The solution was gently mixed and stored on ice for 30 min. The cells were then heat shocked at 42 °C for 2 min and placed on ice briefly (1 min). LB (0.5 mL, no antibiotics) was added to the cells, and the sample was incubated at 37 °C (no agitation) for 1 h. Cells were collected in a microcentrifuge (30 s). If the transformation was to be plated onto LB plates, the cells were resuspended in a small volume of LB medium (0.1 mL), and then spread onto plates containing the appropriate antibiotics. If the transformation was to be plated onto minimal medium plates, the cells was washed once with the same minimal medium. After resuspension in fresh minimal medium (0.1 mL), the cells was spread onto the plates. A sample of competent cells with no DNA added was also carried through the transformation procedure as a control. These

cells were used to check the viability of the competent cells and to verify the absence of growth on selective medium.

**Purification of genomic DNA**

Genomic DNA was purified using a modified method described by Silhavy.[16] A single colony of the strain was inoculated into 100 mL of TB medium (500 mL Erlenmeyer flask). The cells were cultured in a gyratory shaker (37 °C, 250 rpm) for 12 h. Centrifugation (4 000*g*, 5 min, 4 °C) of the culture was followed by resuspension of the cell pellet in 5 mL of buffer [50 mM Tris-HCl (pH 8.0), 50 mM EDTA (pH 8.0)] and storage at -20 °C for 20 min to freeze the suspension. To the frozen cells was added 0.5 mL of 0.25 M Tris-HCl (pH 8.0) that contained 5 mg of lysozyme. The suspension was thawed at room temperature in a water bath with gentle mixing and then stored on ice for 45 min. The sample was then transferred to a Corex tube. After addition of 1 mL of STEP solution [25 mM Tris-HCl (pH 7.4), 200 mM EDTA (pH 8.0), 0.5% SDS (w/v), and proteinase K (1 mg $mL^{-1}$, Sigma), prepared just before use], the mixture was incubated at 50 °C for at least 1 h with gentle, periodic mixing. The solution was then divided into two Corex tubes, and the contents of each tube were extracted with phenol (4 mL). The organic and aqueous layers were separated by centrifugation (1 000*g*, 15 min, room temperature), and the aqueous layer was transferred to a fresh Corex tube. All transfers of the aqueous layer were carried out using wide bore pipette tips to minimize shearing of the genomic DNA. The contents of each tube were extracted again with a mixture of phenol (3 mL) and SEVAG (3 mL). Extractions with phenol/SEVAG were repeated (approximately 6 times) until the aqueous layer was clear.

Genomic DNA was precipitated by addition of 0.1 volume of 3 M NaOAc (pH 5.2), gentle mixing, and addition of 2 volumes of 95% ethanol. Threads of DNA were spooled onto a sealed Pasteur pipette and transferred to a Corex tube that contained 5 mL of 50 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0), and 1 mg of RNase. The mixture was stored at 4 °C overnight to allow the DNA to dissolve completely. The solution was then extracted with SEVAG (5 mL) and centrifuged (1 000 *g*, 15 min, room temperature). The aqueous layer was transferred to a fresh Corex tube and the genomic DNA was precipitated as described above. The threads of DNA were spooled onto a Pasteur pipette and redissolved in 2 mL of 50 mM Tris-HCl (pH 7.5) and 1 mM EDTA (pH 8.0). Genomic DNA was stored at 4 °C.

Alternatively, genomic DNA was purified using a method described by Pitcher.[17] Broth cultures (20 mL) were harvested at the end of the exponential growth phase by centrifugation (1 000*g*, 15 min, room temperature). A small cell pellet was obtained. The cells of Gram-positive species were resuspended in 100 μL of fresh lysozyme (50 mg/mL) in TE buffer and incubated at 37°C for 30 min. The Gram-negative species were resuspended in 100 μl of TE buffer without enzyme treatment and incubated at 37°C for 30 min. Cells were lysed with 0.5 mL 5 M guanidium thiocyanate (Sigma), 100 mM EDTA and 0.5% v/v sarkosyl (GES reagent), which was prepared as follows. Guanidium thiocyanate (60 g), 0.5 M EDTA at pH 8 (20 mL) and deionized water (20 mL) were heated at 65°C with mixing until dissolved. After cooling, 5 mL of 10% v/v sarkosyl were added, the solution was made up to 100 mL with deionized water, filtered through a 0.22-μm membrane and stored at room temperature.

Cell suspensions were vortexed briefly and checked for lysis (clear solution) after 5-10 min. The lysates were cooled on ice and 0.25 mL cold 7.5 M ammonium acetate was added with mixing on ice for 10 min. To this sample, 0.5 mL SEVAG was added, and the solution was mixed thoroughly. After centrifugation in a 1.5 mL Eppendorf tube (25 000*g*, 10 min, room temperature), supernatant fluids were transferred to Eppendorf tubes and 0.54 volumes of cold 2-propanol was added. The tubes were inverted for 1 min to mix the solutions and the fibrous DNA precipitate was deposited by centrifugation (6 500*g*, 20 s, room temperature). Pellets of DNA were washed five times in 70% ethanol and dried at room temperature for 20 min. Genomic DNA was redissolved in 100 μL TE.

**P1-mediated transduction**

Transduction with P1 phage was carried out using a method modified from Miller.[18] P1 phage lysate was prepared by propagation of phage in the donor strain using the following procedure. Serial dilutions of P1 phage stock (0.1 mL, $10^{-1}$ to $10^{-5}$) in LB were prepared in sterile test tubes (13 × 100 mm). An aliquot (0.1 mL, approximately 5 × $10^8$ cells) of an overnight culture of the donor strain was added to each tube. Sterile, molten soft agar (45 °C) was added to each tube. The contents of each tube were mixed and poured immediately onto a pre-warmed (37 °C) L plate, swirling gently to achieve uniform coverage of the plate. After the agar had solidified, the plates were incubated at 37 °C until confluent lysis had occurred (approximately 8 h). Because the multiplicity of infection is critical to phage generation, confluent lysis occurred on only one or two of the plates. L-Broth (4 mL) was added to these plates, which were then stored overnight at 4 °C to allow the phage particles to diffuse into the broth. The L-broth was collected

from the plate and vortexed with several milliliters of $CHCl_3$ to make certain that all of the cells had lysed. The solution was centrifuged (2 000*g*, 5 min, room temperature) to separate the layers. Aqueous phage lysate was stored in 1.5 mL microfuge tubes over several drops of $CHCl_3$ at 4 °C.

Infection of the recipient strain with phage lysate proceeded as follows. Overnight culture (2 mL) of the recipient strain was centrifuged (microfuge, 30 s, 4 °C) and the growth medium discarded. The cells were resuspended in 1 mL of 5 mM $CaCl_2$ and 100 mM $MgSO_4$ and shaken (200 rpm) at 37 °C for 15 min to promote aeration of the cells. In the meantime, 0.1 mL serial dilutions ($10^0$ to $10^{-3}$) of phage lysate in LB were prepared in sterile microfuge tubes. An aliquot (0.1 mL) of aerated recipient cells was added to each of the phage dilutions, the samples were gently mixed and then incubated at 37 °C for 20 min without shaking. Sodium citrate (1 M, 0.2 mL) was added to each sample, and the cells were harvested (microfuge, 30s, room temperature) and resuspended in 0.2 mL of LB containing 100 mM sodium citrate. After incubation at 30 °C for 30 min, cells were again harvested (microfuge, 30 s, room temperature), resuspended in 0.1 mL of growth medium, and plated out onto appropriate agar plates.

**Enzyme assays**

After collected and resuspended in proper resuspension buffer, the cells were disrupted by two passages through a French pressure cell (SLM Aminco) at 16000 psi. Cellular debris was removed from the lysate by centrifugation (48 000*g*, 20 min, 4 °C). Protein was quantified using the Bradford dye-binding procedure.[19] A standard curve

was prepared using bovine serum albumin. The protein assay solution was purchased from Bio-Rad.

**DAHP synthase assay**

DAHP synthase was assayed according to the procedure described by Schoner.[20] Harvested cells were resuspended in 50 mM potassium phosphate (pH 6.5) that contained 10 mM PEP and 0.05 mM $CoCl_2$. The cells were disrupted using a French press as described above. Cellular lysate was diluted in a solution of potassium phosphate (50 mM), PEP (0.5 mM), and 1,3-propanediol (250 mM), pH 7.0. A dilute solution of E4P was first concentrated to 12 mM by rotary evaporation and neutralized with 5 N KOH. Two different solutions were prepared and incubated separately at 37 °C for 5 min. The first solution (1 mL) contained E4P (6 mM), PEP (12 mM), ovalbumin (1 mg/mL), and potassium phosphate (25 mM), pH 7.0. The second solution (0.5 mL) consisted of the diluted lysate. After the two solutions were mixed (time = 0), aliquots (0.15 mL) were removed at timed intervals and quenched with 0.1 mL of 10% trichloroacetic acid (w/v). Precipitated protein was removed by centrifugation, and the DAHP in each sample was quantified using thiobarbituric acid assay[21] as described below.

An aliquot (0.1 mL) of DAHP containing sample was reacted with 0.1 mL of 0.2 M $NaIO_4$ in 8.2 M $H_3PO_4$ at 37 °C for 5 min. The reaction was quenched by addition of 0.8 M $NaAsO_2$ in 0.5 M $Na_2SO_4$ and 0.1 M $H_2SO_4$ (0.5 mL) and vortexed until a dark brown color disappeared. Upon addition of 3 mL of 0.04 M thiobarbituric acid in 0.5 M $Na_2SO_4$ (pH 7), the sample was heated at 100 °C for 15 min. Samples were cooled (2 min), and the pink chromophore was then extracted into distilled cyclohexanone (4 mL).

The aqueous and organic layers were separated by centrifugation (2 000*g*, 15 min, room temperature). The absorbance of the organic layer was recorded at 549 nm ($\varepsilon$ = 68000 L $mol^{-1}$ $cm^{-1}$). One unit of DAHP synthase activity was defined as the formation of 1 μmol of DAHP per min at 37 °C.

**Phosphoenolpyruvate synthase assay**

PEP synthase activity was assayed at 30 °C according to the procedure described by Cooper.[22] The reaction (1 mL) contained 100 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 10 mM ATP, 1.25 mM sodium pyruvate and was initiated by addition of cell-free lysate. Aliquots (100 μl) of reaction mixture were removed at 1 min intervals and immediately added to a microcentrifuge tube containing 0.33 mL of a 0.1% aqueous solution of 2, 4-dinitrophenylhydrazine and 0.9 mL $H_2O$. The resulting mixture was incubated at 30 °C for 10 min. Following addition of 1.67 mL of 10% (w/v) NaOH, the mixture was further incubated at 30 °C for 10 min. The disappearance of pyruvate was quantified by measuring the absorbance at 445 nm. A molar extinction coefficient of 18,000 L $mol^{-1}$ $cm^{-1}$ was used to quantify pyruvate. One unit of PEP synthase activity was defined as the amount of enzyme that catalyzed the consumption of 1 μmol of pyruvate per min at 30 °C.

**Glucokinase assay**

Glucokinase was assayed by coupling with glucose-6-phosphate dehydrogenase according to the method of Fraenkel.[23] The assay solution contained 50 mM Tris-HCl (pH 7.6), 0.5 mM glucose, 2 mM $MgCl_2$, 2 mM ATP, 1 μg of glucose 6-phosphate

dehydrogenase (3 units), cell-free lysate and was initialized by addition of 0.4 mM NADP. The formation rate of glucose 6-phosphate was equal to the formation rate of NADPH which was measured by the increase in absorbance at 340 nm ($\varepsilon$ = 6220 L $mol^{-1}$ $cm^{-1}$). One unit of glucokinase activity was defined as the formation of 1 μmol of glucose 6-phosphate per min at 25 °C.

**Transketolase assay**

Transketolase was assayed using a coupled enzyme procedure described by Paoletti.[24] The assay solution (1 mL) contained triethanolamine buffer (150 mM, pH 7.6), $MgCl_2$ (5 mM), thiamine pyrophosphate (0.1 mM), NADP (0.4 mM), β-hydroxypyruvate (0.4 mM), D-erythrose 4-phosphate (0.1 mM), glucose 6-phosphate dehydrogenase (3 units), and phosphoglucose isomerase (10 units). The solution was incubated at room temperature and the absorbance at 340 nm was monitored for several minutes. After all the unreacted D-glucose-6-phosphate from the D-erythrose 4-phosphate synthesis had reacted, an aliquot of transketolase solution was added and the reaction monitored at 340 nm for 10-20 min. One unit of transketolase activity was defined as the formation of 1 μmol of NADPH ($\varepsilon$ = 6220 L $mol^{-1}$ $cm^{-1}$) per min.

## Chapter 2

### *E. coli* JY1 (KL3 Δ*ptsHIcrr*)

*E. coli* JY1 (KL3 Δ*ptsHIcrr*) was made from *E. coli* KL3 by P1 transduction using *E. coli* UE79 as the donor strain. Colonies were plated on M9 maltose plates containing fosfomycin with aromatic and serine supplementation. The resulting colonies

were screened for loss of PTS activity by checking the growth on the following plates: no growth on M9 glucose plates with aromatic, serine supplementation; no growth on M9 xylose plates with aromatic and serine supplementation; growth on M9 xylose plates with aromatic, serine and cAMP (1 mM final concentration on the plate) supplementation; no growth on M9 mannitol plates with aromatic and serine supplementation.

### *E. coli* JY1.2 (KL3 Δ*ptsHIcrr galP*$^+$)

Plasmid pRC55B containing *serA* gene was transformed into E. *coli* JY1. The resulting *E. coli* JY1/pRC55B was used to inoculate 1 L of fermentation medium containing 20 g glucose.[14] Doubling times for growth under fermentor-controlled conditions were determined[18] by dilution of fermentation broth (1:100) followed by measurement of absorption at 600 nm ($OD_{600}$). After cultivation under fermentor-controlled conditions for 3.5 days, the culture's doubling time decreased from 12 h to 3.7 h. When the culture reached $OD_{600}$ = 20, fermentor broth (900 mL) was discarded and replaced with fresh fermentation medium containing 20 g glucose. Repetition of this dilution process two more times led to a culture with a doubling time of 2.7 h. Serial dilutions of the final culture were spread onto M9 glucose plates with aromatics supplementation and incubated at 37 °C. Single colonies were subsequently obtained after growth on M9 glucose plates. These colonies were screened for growth on M9 maltose plates containing fosfomycin with aromatic supplementation and an absence of growth on M9 mannitol plates with aromatic supplementation. Colonies possessing the proper phenotype were also screened on M9 glucose plates with aromatic supplementation for the most rapid growth. One colony with most rapid growth was

inoculated into 5 mL LB culture and grown at 37°C for 12 h. Cultures were diluted (1:10000) in LB, and two more cycles of growth at 37 °C for 12 h were carried out to promote plasmid loss from the cells. Serial dilutions of the final culture were spread onto LB plates and the resulting colonies were screened for resistance to Cm. A $Cm^S$ colony designated as *E. coli* JY1.2 was isolated that retained the phenotype for the absence of PTS-mediated glucose transport as characterized by growth on M9 maltose plates containing fosfomycin with aromatic and serine supplementation and an absence of growth on M9 mannitol plates with aromatic and serine supplementation.

***E. coli* JY1.2 *galP::Tn10***

*E. coli* JY1.2 was made *galP::Tn10* by P1 transduction from *E. coli* JM2053. Colonies were plated onto LB plates containing Tc to select for colonies which had been transduced by *galP::Tn10.* The resulting *E. coli* JY1.2 *galP::Tn10* didn't grow on M9 glucose plates with aromatic and serine supplementation.

***E. coli* JY1.3 (KL3 Δ*ptsHIcrr galP*$^+$)**

*E. coli* JY1.2/pRC55B was used to inoculate a 5 mL liquid M9 glucose medium with aromatic supplementation and grown at 37°C until the culture turned turbid. Cultures were diluted (1:10000) in M9, and four more cycles of growth at 37 °C were carried out to obtain fast growing mutants of *E. coli* JY1.2/pRC55B. Serial dilutions of the final culture were spread onto M9 glucose plates with aromatic supplementation and incubated at 37 °C for 24 h. The biggest colony on the M9 glucose plate was inoculated into 5 mL LB culture and grown at 37°C for 12 h. Cultures were diluted (1:10000) in LB,

and two more cycles of growth at 37 °C for 12 h were carried out to promote plasmid loss from the cells. Serial dilutions of the final culture were spread onto LB plates and the resulting colonies were screened for resistance to Cm. A $Cm^S$ colony designated as *E. coli* JY1.3 was isolated that retained the phenotype for the absence of PTS-mediated glucose transport as characterized by growth on M9 maltose plates containing fosfomycin with aromatic and serine supplementation and an absence of growth on M9 mannitol plates with aromatic and serine supplementation.

### *E. coli* JY1.3 *galP::Tn10*

*E. coli* JY1.3 was made *galP::Tn10* by P1 transduction from *E. coli* JM2053. Colonies were plated onto LB plates containing Tc to select for colonies which had been transduced by *galP::Tn10*. The resulting *E. coli* JY1.3 *galP::Tn10* didn't grow on M9 glucose plates with aromatic and serine supplementation.

### Plasmid pJY1.131

Digestion of pKL4.66B[2] with *Bam*HI and *Sal*I afforded a 1.3-kb fragment of DNA that encoded $aroF^{FBR}$. Localization of the $aroF^{FBR}$ fragment in pJF118EH which had been previously digested with *Bam*HI and *Sal*I resulted in pJY1.131. Transcription of the $aroF^{FBR}$ gene in pJY1.131 proceeds in the opposite orientation relative to the vector-encoded $P_{tac}$ sequence.

### Plasmid pJY1.143A

The *ppsA* gene was originally amplified from RB791 genomic DNA using the following primers: 5'-AACTGCAGCGATCCAGTTTCATCTCTTGT and 5'-AACTGCAGTTAT TTCTTCAGTTCAGCCAG. *Pst*I restriction sequences (underlined nucleotides) were included to facilitate cloning. Localization of the resulting 3.0-kb *ppsA* fragment into the *Pst*I site of pSU19 afforded pKL1.87A. The *ppsA* ORF was subsequently amplified from pKL1.87A using the following primers: 5'-GGAATTCTTATTTCTTCAGTTCAGCCAG and 5'-GGAATTCATGTCCAACAATGGCTCGT. Insertion of the resulting *ppsA* ORF-encoding fragment into the *Eco*RI site of pJY1.131 afforded the 9.0-kb plasmid pJY1.143A in which the *ppsA* ORF is transcribed from the vector-encoded $P_{tac}$ sequence that immediately precedes the gene.

**Plasmid pJY1.207A**

A 0.2-kb sequence of DNA encoding the three operator binding sites associated with the *aroF* gene, designated $P_{aroF}$, was amplified from pMF63A using the following primers: 5'-GCGGATCCGAATTCAAAGGGAGTGTA and 5'-GCGGATCCCCTCAGCGAGG ATGACGT. Localization of the $P_{aroF}$ fragment of DNA into the *Bam*HI site of pJY1.143A afforded pJY1.207A.

**Plasmid pJY1.211A**

Ligation of a 1.9-kb *Dra*I/*Eco*RV *serA*-encoding fragment of DNA obtained from pD2625 into the *Sma*I site of pJY1.207A resulted in isolation of pJY1.211A in which

transcription of *serA* proceeds in the opposite orientation relative to the vector-encoded $P_{tac}$ sequence.

**Plasmid pJY1.216A**

Following digestion of pMF51A with *Bam*HI, the resulting 2.2-kb *tktA* fragment was treated with Klenow fragment and ligated to pJY1.211A which had previously been incubated with *Hin*dIII followed by Klenow fragment. The 13.3-kb plasmid pJY1.216A was isolated in which transcription of the *tktA* gene proceeds in the orientation opposite to transcription from the vector-encoded $P_{tac}$.

**Plasmid pKL7.115A**

Plasmid pKL7.115A is a pSU18-derived plasmid that encodes $aroF^{FBR}$, $P_{aroF}$, *serA*, $P_{tac}$*glf*, *tktA*, and $Cm^R$. Construction of pKL7.115A began with pKD11.291A, a 5.6-kb plasmid that was described previously.[2] A 2.2-kb $P_{tac}$*glf*-encoding fragment was liberated from pTC325 by digestion with *Bam*HI, *Hin*dIII and *Xba*I. Insertion of the $P_{tac}$*glf* fragment into *Sal*I site of pKD11.291A yielded pKL6.239A. Following digestion of pMF51A with *Bam*HI, the resulting 2.2-kb *tktA*-encoding fragment was inserted into *Hin*dIII site of pKL6.239A to afford the 10.0-kb plasmid pKL7.115A.

**Plasmid pJY2.182**

The 5.0-kb plasmid was created by ligation of a 2.2-kb *tktA* fragment obtained by *Bam*HI digestion of pMF51A into the *Bam*HI site of p34H.

**Plasmid pJY2.183A**

Plasmid pJY2.183A is a pSU18-derived plasmid that encodes *aroF*$^{FBR}$, $P_{aroF}$, *serA*, $P_{tac}$*glfglk*, *tktA*, and Cm$^{R}$. Construction of pJY2.183A began with pKD11.291A. A 3.9-kb $P_{tac}$*glfglk*-encoding fragment was liberated from pTC325 by digestion with *Bam*HI and *Xba*I. Insertion of the $P_{tac}$*glfglk* fragment into *Sal*I site of pKD11.291A yielded pKL7.85A. Following digestion of pJY2.182 with *Sac*I, the resulting 2.2-kb *tktA*-encoding fragment was inserted into *Sac*I site of pKL7.85A to afford the 11.7-kb plasmid pJY2.183A.

**Plasmid pJY3.29B**

Plasmid pJY3.29B is a pSU18-derived plasmid that encodes *aroF*$^{FBR}$, $P_{aroF}$, *serA*, *glk*, *tktA*, and Cm$^{R}$. Construction of pJY3.29B began with pKD11.291A. A 2.2-kb *glk*-encoding fragment was liberated from pTC325 by digestion with *Hin*dIII. Insertion of the *glk* fragment into *Sal*I site of pKD11.291A yielded pJY3.21B. Following digestion of pJY2.182 with *Sac*I, the resulting 2.2-kb *tktA*-encoding fragment was inserted into *Sac*I site of pJY3.21B to afford the 10.0-kb plasmid pJY3.29B.

**Plasmid pRC55B**

The 3.9-kb plasmid was created by ligation of a 1.6-kb *serA* fragment obtained by PCR from *E. coli* RB791 genomic DNA into the *Sma*I site of pSU18.

## Chapter 3

### Isolation of total RNA

Total RNA was isolated with hot phenol and used to prepare [$\alpha$-$^{33}$P]dCTP-labeled cDNA as described by Tao.[25] All the solutions are 0.1% diethyl pyrocarbonate (DEPC, Aldrich) treated for 24 h at 37 °C and autoclaved for 25 min to avoid RNase contamination. All the glass bottles used to store solutions were prepared by filling the bottles with 0.1% DEPC solutions at 37 °C for 24 h followed by discarding 0.1% DEPC solutions and autoclaving the empty bottles. Phenol (Ultrapure, Invitrogen) was equilibrated with an equal volume of 50 mM NaOAc (2.07 g in 500 mL water and adjust pH to 4.8 using HOAc) three times in a small bottle. Before use, 0.1% 8-hydroxyquinoline was added to this acidic phenol and stored at 4 °C for up to one week. Unless specified, all the enzymes were purchased from Invitrogen.

A sample of the *E. coli* culture from the fermentation (0.1 to 0.2 mL) was rapidly transferred to a 17 × 100 mm sterilized test tube containing 4 mL of hot lysis buffer (1% sodium dodecyl sulfate, 30 mM sodium acetate, 3 mM EDTA; pH 5.0) in a boiling water bath close to the fermentor and held for 5 min with intermittent mixing. The resulting lysates were extracted with 3 mL acidic phenol (prewarmed to 65 °C) with vigorous intermittent mixing for 3 min. The use of acidic phenol was essential to minimize genomic DNA contamination of the RNA sample. After cooling the sample on ice for 3 min, the sample was centrifuged (12 000*g*, 10 min, 4 °C). The aqueous layer was extracted a second time with 3 mL acidic phenol (prewarmed to 65 °C) followed by extraction with equal volumes (1.5 mL) of phenol and chloroform at room temperature.

After the final extraction with 3 mL of chloroform at room temperature, the aqueous layer was transferred to a fresh tube. Care must be taken to avoid transferring any of the organic-phase along with the aqueous-phase. The aqueous layer was divided equally between two fresh tubes and one-tenth volume of 3 M sodium acetate (pH 5.2, ice cold) was added followed by 2.5 volumes of 100% ethanol (-20 °C). The sample was kept at –20 °C overnight. Total RNA (two tubes) was pelleted by centrifugation (12 000*g*, 20 min, 4 °C), washed carefully by addition of 70% ethanol (ice cold), and then pelleted again by centrifugation (12 000*g*, 20 min, 4 °C). The RNA pellet (two tubes) was washed again by addition of 70% ethanol (ice cold) followed by centrifugation (12 000*g*, 20 min, 4 °C). The RNA pellet (two tubes) at the bottom of the tube was marked outside the tube with a pen and then dried at room temperature under air for 2-3 h. When the RNA pellet (two tubes) appeared clear or translucent, RNA was dissolved in 90 μL of sterile, RNase-free water by repeatedly pipeting. This time the RNA solutions from two tubes were combined (total volume was about 180 μL) and stored on ice. Contaminating DNA was removed from RNA by hydrolysis with DNaseI. RNA (total volume approximately 180 μL), 20 μL DNaseI reaction buffer (10X, provided by the manufacturer), DNaseI (8 μL, 200 U) and RNase inhibitor (2 μL, 80 U) was incubated at room temperature for 25 min. After the reaction, Qiagen Rneasy Mini Kit was used to purify the RNA according to the procedure provided by the manufacturer. The purified RNA sample (approximately 160 μL) was quantified by measuring the absorbance at 260 nm (usually 10 μL of RNA in 1 mL water). The RNA concentration was calculated based on the fact that the absorbance at 260 nm of a 40 μg $mL^{-1}$ of RNA is 1.0. The integrity of the RNA and the amount of genomic DNA contamination was checked by

agarose gel electrophoresis. A non-denaturing agarose gel was used so that any genomic DNA contamination could be easily observed. The TAE buffer used to run the gel was pretreated with 0.1% DEPC as described previously. The RNA sample (10 μL) was loaded with 1 μL RNase inhibitor (Invitrogen, 40 U) and stained with ethidium bromide. The 23S and 16S ribosomal RNA bands should be clearly visible at about 2:1 ratio (23S: 16S) of staining intensity. If genomic DNA was present in the RNA sample, it would be seen as high molecular weight-staining material. The final RNA sample was kept at –80 °C after addition of one-tenth volume 3 M sodium acetate (pH 5.2, ice cold) and 2.5 volumes of 100% ethanol (-20 °C).

**Synthesis of radiolabeled cDNA**

A solid mixture of primers homologous to the 3' ends of the predicted 4,290 open reading frame (ORFs) in *E. coli* was obtained from Sigma-Genosys (Panorama *E. coli* cDNA labeling primers) and was resuspended in 40 μL sterile RNase-free water before use. The *E. coli* RNA (5 μg) was pelleted from stored RNA sample (–80 °C) by centrifugation (microfuge, 20 min, 4 °C), washed carefully by addition of 20 μL 70% ethanol (ice cold), and then pelleted by centrifugation (microfuge, 20 min, 4 °C). The RNA was dissolved in 8.5 μL of sterile RNase-free water, mixed with 8 μl of the prepared *E. coli* cDNA labeling primers, and incubated in a heat block at 90°C for 2 min to denature the RNA. Then the *E. coli* cDNA labeling primers were annealed to the RNA template by removing the block from the heating device and placing on the work bench to cool the sample slowly to 42°C. To this sample 2 μL deoxynucleoside triphosphate mix (lacking dCTP, 10 mM each), 3 μL DTT (0.1 M), 7 μL reverse transcriptase buffer,

1.5 μL dCTP (0.2 mM), 2 μL Superscript II Rnase H$^-$ reverse transcriptase (400 U), 1 μL RNase inhibitor (40 U), and 2 μL [α-$^{33}$P]dCTP (20 μ Ci; specific activity, 2000-3000 μCi/mmol; New England Nuclear) were added to initiate cDNA synthesis (final volume 35 μL). The reaction mixture was incubated at 42 °C for 1 h. To this sample 1.5 μL EDTA (0.5 M, pH 8.0) and 1.5 μL NaOH (5 M) was added to degrade the RNA at 65 °C for 30 min. The sample was neutralized by adding 4 μL Tris-HCl (1 M, pH 8.0) and 3.5 μL HCl (2 M) followed by addition of 54.5 μL $H_2O$ to make the total volume 100 μL. An aliquot (2 μL) of this sample was withdrawn to serve as a control for measuring total radioactivity by scintillation counter. In the remaining sample, radioactive cDNA was separated from other small molecules by gel filtration using a Micro Bio-spin P-30 Tris Chromatography Column (Bio-rad) according to the procedure provided by the manufacturer. An aliquot (2 μL) of the purified cDNA was withdrawn to measure the radioactivity of the synthesized cDNA by scintillation counter. The remaining purified cDNA was ready for hybridization. Approximately 50-70% of total radioactive label was incorporated into synthesized cDNA.

To measure the radioactivity by scintillation counter, 2 μL sample was directly added to a small filter paper. The filter paper was put into a vial filled with 10 mL 3a20™ complete counting cocktail (Research Products International). After shaking the vial vigorously, the radioactivity of the sample was recorded in a 1414 Winspectral liquid scintillation counter [PerkinElmer LS (WALLAC)].

**Hybridization**

After preparing the radioactively labeled cDNA, the next step was to perform a hybridization to the Panorama gene array (Sigma-Genosys) containing duplicate, UV-cross-linked arrays of 4,290 ORF-specific DNAs. All blots (filters) used in these experiments were from the same manufacturing lot. Each blot (filter) was stripped and used a maximum of three times. A total of 24 hybridizations (48 arrays) were analyzed. The hybridizations were performed in roller bottles at 65 °C in a hybridization incubator (Fisher Biotech), where minimal volumes of hybridization solutions are employed. Before use, the temperature control in the hybridization incubator was calibrated according to the procedure provided by the manufacturer.

The hybridization and washing steps were carried out as described by Tao.[25] The hybridization incubator was prewarmed to 65 °C. The roller bottles were filled with water and placed into hybridization incubator to check if they were leaking. Sheared salmon sperm DNA (Invitrogen, 100 ug/mL), which is a blocking agent to reduce the background in hybridization experiments,[15] was boiled on 95 °C water bath for 15 min followed by rapid cooling on ice to denature the DNA. The 10 mL hybridization solution were prepared by mixing 1 mL 20% sodium dodecyl sulfate (SDS), 2.5 mL 20X SSPE (3.6 M NaCl, 0.2 M sodium phosphate, 20 mM EDTA; pH 7.4), 6.2 mL distilled water, 0.2 mL 50X Denhardt's solution (USB, prewarmed to 65 °C) and 0.1 mL sheared salmon sperm DNA.[25] The blot was rinsed twice in 400 mL 2X SSPE (0.36 M NaCl, 20 mM sodium phosphate, 2 mM EDTA; pH 7.4) for 10 min in a big container. A hybridization nylon mesh (Hybaid) was cut into certain size (24.5 cm × 12.5 cm) and used to wrap the blot. The wrapped blot was inserted into the roller bottle filled with 50 mL 2X SSPE (prewarmed to 65 °C). After washing the blot twice with 50 mL 2X SSPE (prewarmed to

65 °C) inside the roller bottle, 5 mL prepared hybridization solution (prewarmed to 65 °C) was added to the roller bottle to initiate prehybridization at 65 °C. After 6 h of prehybridization, the hybridization solution inside the roller bottle was discarded and the labeled cDNA (boiled on 95 °C water bath for 15 min followed by rapid cooling on ice for 5 min to denature) in a fresh 4 mL hybridization solution was added. After hybridization for 16 h at 65 °C, the blot was washed three times (5 min per wash) by inverting the roller bottle by hand at room temperature and three times at 65 °C (25 min per wash) in the hybridization incubator with 80 mL 0.5X SSPE containing 0.2% sodium dodecyl sulfate. The wet blot was wrapped in Mylar film (1.5 μm, Persuation Marketing Specialists) and was exposed to a Low Energy Storage Phosphor Screen (Molecular Dynamics) for 20 h. The imaging screen was scanned at a 50 μm pixel size rather than a 100 μm or 200 μm pixel size, for greater resolution of spots, using a STORM 820 Phosphoimager (Molecular Dynamics) in the Biochemistry Department of Michigan State University. Each image file was about 50 MB and was backed up immediately. The blot was stripped at 100 °C with 1% SDS in Tris-EDTA buffer according to the procedure specified by the manufacturer. The same blot was hybridized, stripped, and rehybridized for 3 times. The stripped blot was wrapped in a plastic food wrap and was kept at –20 °C until ready to use.

**Quantitation of cDNA as a relative measure of gene expression**

The image files were analyzed by determining the pixel density (intensity) for each spot in the array by using Array Vision software (version 6.1; Imaging Research Inc.). A grid of individual ellipses corresponding to each of the DNA spots on the blots

was laid down on the image to designate each spot to be quantified. Background was subtracted automatically by the Array Vision software using the "surrounding entire template" method. The intensities for each spot were exported from Array Vision software into a Microsoft Excel spreadsheet. The sum of values for all 4,290 ORFs in each array was normalized to 100 million to allow comparisons of gene expression between experiments. For example, intensity value of 10000 corresponded to 0.01% of total cDNA hybridized to the array. For each condition investigated (with or without IPTG addition, different times during fermentation; 8 different conditions), expression levels for 6 arrays (3 filters with duplicate arrays) were measured. Results for every time point represent an average of six determinations (three hybridizations each, two arrays per filter). A threshold value for reliable detection was established as 0.002% of total bound cDNA (intensity value 2000).[25] An average expression level for all 8 conditions was calculated for each gene using Microsoft Excel. The 347 genes with an average expression intensity below 2000 were not included in further statistical analysis.

**Statistical analysis**

To address the issue of reproducibility of the gene expression data, the coefficient of variation (standard deviation expressed as a percentage of the mean) was used as a standard measure of gene expression data quality described by Tao.[25] The standard deviation is a measure of how spread out a distribution in a sample is. The formula for computing standard deviation ($\sigma$) in a sample using Microsoft Excel (function STDEV) is $\sigma = \sqrt{\frac{n\sum x^2 - \left(\sum x\right)^2}{n(n-1)}}$ where n is the number of scores in a sample. The coefficient of

variation for each gene under 8 different conditions was calculated (n = 6) and the average coefficient of variation of total 3943 genes under each condition was calculated.

Student's *t* test is one of the most commonly used techniques for testing a hypothesis on the basis of a difference between sample means. Explained in layman's terms, the *t* test determines a probability *p* that two populations are the same (null hypothesis) with respect to the variable tested. Some researchers suggested that if the difference is in the predicted direction, only one half (one "tail") of the probability distribution is considered and thus divide the standard *p*-level reported with a *t*-test (a "two-tailed" probability) by two. Others, however, suggested that the standard, two-tailed *t*-test probability should always be reported. The only assumption in the student *t* test is that the population follows a normal distribution (small samples have more scatter and follow what is called a *t* distribution). Because log values of expression data have previously been shown to approximate a normal distribution,[25] log values of expression data were generated using Microsoft Excel and the significance of gene comparisons between two samples (with IPTG treatment and without IPTG treatment) was evaluated by computing the probability values for the null hypothesis using two-tailed student *t* test in Microsoft Excel (function TTEST). The *t* test can be performed knowing just the means, standard deviation, and number of data points for the two samples. The two samples *t* test yields an intermediate statistic *t*, in which $t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{m} + \frac{\sigma_y^2}{n}}}$. Note that the numerator of the formula is the difference between means of two samples (x-bar and y-bar is the sample mean). The denominator is a measurement of experimental error in the two samples combined ($\sigma_x$ and $\sigma_y$ is the sample standard deviation, m and n is the number

of scores in the sample). The higher the value of $t$, the greater the confidence that there is a difference. A precise probability value $p$ can be attached to that confidence using the $t$ distribution table originally worked out by W.S. Gossett.[26] The following formula was used to approximate the degrees of freedom used in the $t$ distribution table:

$$df = \frac{\left(\frac{\sigma_x^2}{m} + \frac{\sigma_y^2}{n}\right)^2}{\frac{\left(\sigma_x^2/m\right)^2}{m-1} + \frac{\left(\sigma_y^2/n\right)}{n-1}}.$$

Cluster analysis was used to compare expression data for each condition (with or without IPTG addition, different times during fermentation). Clusting analysis was performed using hierarchical clusting methods by Cluster and TreeView software (download from http://rana.lbl.gov/ in Esien's lab) as described by Eisen.[27] The gene (condition) similarity used was a form of correlation coefficient. For any two conditions X and Y observed over a series of N genes, a similarity score was computated as follows: $S(X,Y) = \frac{1}{N}\sum_{i=1,N}\left(\frac{X_i - \overline{X}}{\sigma_X}\right)\left(\frac{Y_i - \overline{Y}}{\sigma_Y}\right)$ where X-bar and Y-bar is the mean of observations and $\sigma_X$ and $\sigma_Y$ is the standard deviation of the observations. For any sets of n conditions, an upper-diagonal similarity matrix was computed by using the similarity formula described above, which contained similarity scores for all pairs of conditions. The matrix was scanned to identify the highest value (respresenting the most similar pair of conditions). A node was created joining these two conditions, and a gene expression profile is computed for the node by averaging observation for the joined elements. The similarity matrix was updated with this new node replacing the two joined elements, and the process was repeated n-1 times until only a single element remains.

**Medium for PCR mutagenesis of *aroG***

The selection medium for PCR mutagenesis of *aroG* gene consisted of $KH_2PO_4$ (3 g/L), $Na_2HPO_4$ (6 g/L), $NH_4Cl$ (1 g/L), NaCl (0.5 g/L), $MgSO_4$ (1 mM), glucose (4 g/L), thiamine (6 mg/L), leucine (25 mg/L), nicotinic acid (6 mg/L) and phenylalanine (3.3 g/L).

The SOC medium used for transformation of electrocompetent cells was prepared as the following. Tryptone (2 g), yeast extract (0.5 g), NaCl (1 mL 1 M NaCl), KCl (0.25 mL 1 M KCl) was added to 97 mL of distilled water. After dissolving and autoclaving, 1 mL 1 M $MgCl_2$ and 1 mL 2 M glucose was added to the solution. The complete medium was filtered through a 0.22 μm filter.

**Plasmid pJY6.119A**

This 6.3-kb plasmid was constructed by ligation the open reading frame (ORF) of *aroG* into the *Sma*I/*Bam*HI site of pJF118EH. The *aroG* ORF was amplified from *E. coli* RB791 genomic DNA using the following primers: 5'-CGGGATCCTTACCCCGCGACGCGCTTT and 5'-TCCCCCGGGATGAATTATCAGAACGACGAT. Localization of the resulting 1.0-kb fragment into pJF118EH afforded pJY6.119A.

**Plasmid pJY6.149 (#1-#13)**

The *aroG* ORF from plasmid pJY6.119A was subject to PCR mutagenesis using the following primers: 5'-CGGAATTCATGAATTATCAGAACGACGAT and 5'-

CGGGATCCTTACCCGCGACGCGCTTT. The PCR product was purified and digested with *Eco*RI/*Bam*HI. The resulting DNA was ligated to pJF118EH, which has been digested with *Eco*RI/*Bam*HI. The resulting ligation mixture was introduced by transformation into CB734 electrocompetent cells and the resulting transformants were selected at 37 °C on the minimal selection plate describe above. The 13 plasmids pJY6.149 (#1-#13) were isolated from the colonies that grew on the selection plate.

**Plasmid pJY6.186**

A 12.2-kb DNA fragment was amplified by Long PCR (described below) from plasmid pJY1.216A using the following primers: 5'-GAAGATCTGATCTGAACGGGCAGCTGAC and 5'-GAAGATCTCGATCCTGTTTATGCTCGTTTG. The PCR product was digested with *Bgl*II and religated by itself to afford 12.2-kb plasmid pJY6.186.

**Plasmid pJY6.231A**

This 13.2-kb plasmid was constructed by ligation the ORF of *aroG*$^{FBR}$ (digested with *Bam*HI) into the *Bgl*II site of pJY6.186. The *aroG*$^{FBR}$ ORF was amplified by PCR from plasmid pJY6.149#4 using the following primers: 5'-CGGGATCCATGAATTATCAGAACGACGAT and 5'-CGGGATCCTTACCCGCGACGCGCTTT. The ORF of *aroG*$^{FBR}$ is transcribed in the same orientations as the *aroF* promoter in pJY6.186.

**PCR mutagenesis of *aroG* gene**

The protocol for mutagenic PCR is derived from standard PCR conditions as described by Cadwell.[28] Plasmid pJY6.119A was digested with *Sma*I/*Bam*HI and the *aroG* ORF DNA was purified by gel electrophoresis. Mutagenic PCR buffer (10X) contained 70 mM $MgCl_2$, 500 mM Kcl, 100 mM Tris (pH 8.3), and 0.1% (w/v) gelatin. Deoxynucleoside triphosphate mix (10X) contained 2 mM dGTP, 2 mM dATP, 10 mM dCTP, and 10 mM dTTP. The mutagenic PCR reaction was prepared by mixing 10 μL 10X mutagenic PCR buffer, 10 μL 10X dNTP mix prepared above, 14 μL 50 mM $MgCl_2$, 30 pmoles of each primer, 20 ng of input DNA, and an amount of water that brought the total volume to 88 μl. To this sample was added 10 μL 5 mM $MnCl_2$ (make sure a precipitate was not formed) and 2 μl Taq polymerase (5 Unit), bringing the final volume to 100 μL. The PCR reaction was performed in a thermal cycler (Eppendorf, Mastercycler Gradient). The sample was incubated for 30 cycles of 94°C for 45 s, 45°C for 1 min, and 72°C for 45 s. A "hot start" procedure or a prolonged extension time at the end of the last cycle was not employed.

**Preparation of electrocompetent cells (*E. coli*)**

*E. coli* CB734 was prepared for electroporation with the following procedure. *E. coli* strain CB734 was initially grown from a single colony in 5 mL LB/Cm and shaken for 16 h at 37°C . From the 5 mL culture, 2 mL was added to 500 mL 2 × YT medium (Bacto-tryptone 16 g, yeast extract 5 g, NaCl 5 g per liter) which was grown until reaching an absorbance of 0.6 - 0.8 at 600 nm. The cells were chilled to 4°C then harvested by centrifugation (3 000*g*, 5 min, 4 °C). The supernatant was discarded, and the cell pellet was resuspended in 500 mL ice cold water. The cells were harvested by

centrifugation (3 000*g*, 10 min, 4 °C) then resuspended in 250 mL ices cold water. Following harvesting by the same conditions, the cells were resuspended in 100 mL ice cold 10% glycerol solution (Sterile filtered, not autoclaved). The cells were once again harvested by the same conditions and resuspended in 1.5 mL ice cold 10% glycerol solution. 0.1 mL aliquots were fast frozen in liquid nitrogen and store at –80 °C.

For transformation, 40 μL of cells were combined with plasmid DNA followed by incubation for 2 min at 0 °C. The cells were then transferred to a pre-chilled electroporation cuvette with a 0.2 cm gap width. The cells were electroporated at 2.5 kV, 25 μF, and 200 Ohms using Gene Pulser II (Bio-rad). Following electroporation, 1 mL room temperature SOC medium was immediately added to the cuvette and incubated at 37 °C for 1 h. The entire cell solution was plated with the appropriate selection plate.

**Long PCR**

The long PCR (12.2 kb) from plasmid pJY1.216A was performed according to the modified protocol of Cheng.[29] Efficient Long PCR results from the use of two polymerases: a non-proofreading polymerase is the main polymerase in the reaction, and a proofreading polymerase is present at a lower concentration. The Tth DNA polymerase (Epicentre Technologies) was used as the main component and Vent DNA polymerase (Invitrogen) as the fractional-component polymerase. Each reaction contained 5 U Tth DNA polymerase, 0.2 U Vent DNA polymerase, 1X Long PCR buffer [(85 mM KOAc, 25 mM tricine (pH 8.7), 8% glycerol, 1% DMSO, 1.2 mM $Mg(OAc)_2$], 0.2 mM each dNTP, 50 pmoles each primer and 40 ng template DNA. The target sequences were amplified through 15 cycles of 94 °C for 10 s, 64 °C for 1 min and 72 °C for12 min 18 s,

followed by another 15 cycles of 94 °C for 10 s, 64 °C for 1 min and 72 °C for12 min 18 s + 30 s/cycle. A hot start method was employed for Long PCR. The reaction was divided into two parts: a template/primer fraction which was 4/5 of the reaction volume, and a polymerase fraction which constituted the remaining 1/5 of the reaction. Each fraction was 1X buffer concentration. The polymerase fraction contained only polymerase, buffer and water. All other components were included in the template/primer fraction. The template/primer fraction was incubated in the PCR machine to 94 °C for 10 s to denature. After denaturing, an 80 °C step was used for adding the polymerase fraction. The following guidelines were used for picking primers for Long PCR: primers are 20 to 23 bases in length; G + C = 12 bases; A + T = 8 to 11 bases; ideal Tm = 60 to 68 °C; avoid primer hairpins; avoid primers with 3' complementarity (results in primer-dimers). The primers were designed with the aid of Oligo software.

## Chapter 4

### Silicon oil centrifugation

Cells were separated from the medium and inactivated by silica oil centrifugation with the further working up procedure described by Ebbighausen.[30] Cells from fermentation experiments were harvested by centrifugation (microfuge, 30 s, 4 °C) followed by one wash cycle in a buffer contain 50 mM Tris-HCl, pH 7.5, 10 mM NaCl, 20 mM KCl. The final pellet was suspended in the same buffer at a cell density of about $10^9$ cells/mL ($OD_{600}$ = 5-10). An aliquot (200-400 μL) was transferred to a 1.5 mL

microfuge tube containing a layer of 260 μL silicone oil (AR200, Fluka) floating on a layer of 120 μL 20% perchloric acid. Cells (800 μL) from shake-flask experiments were directly transferred to the above microfuge tube. The tube was immediately centrifuged (microfuge, 1 min, 4 °C) followed by removing the supernatant and silicone oil layer. After homogenization the sediment by sonication (six 30 s bursts and 30 s intermittent cooling in ice), the perchloric extracts were neutralized with a solution containing 5 M KOH/1 M triethanolamine and chilled on ice for 10 min to precipitate the $KClO_4$ formed on neutralization. Together with the denatured protein and residual oil the precipitated salt was then sedimented by centrifugation (microfuge, 5 min, room temperature) and the supernatant was used for shikimic acid analysis.

**Shikimic acid quantitation**

A recently reported chromogenic assay that facilitates rapid detection and quantitation of shikimic acid[31] was used to analyze the above supernatant. Supernatant (10 μL – 50 μL) was added with water to make final volume 250 μL. Then 250 μL of periodate reagent (0.5% periodic acid and 0.5% sodium meta-periodate) was added. After 30 minutes at 37 °C, the reactions were quenched with 0.3 mL of 1 N NaOH followed by 0.2 mL of 0.056 M sodium sulfite. Absorbance of the solution at 382 nm was used to quantitate the amount of shikimic acid in the solution. A standard concentration curve was determined for shikimic acid using solutions of authentic, purified shikimic acid.

**Shake flask experiment of *E. coli* SP1.4**

*E. coli* SP1.4 cells were grown in 100 mL LB medium for 12 h, washed with 100 mL M9 and resuspended in 100 mL M9 minimal medium with glucose and serine addition. After cultivation of cells for 36 h at 37 °C, the supernatant was analyzed by $^{1}H$ NMR.

**Mini-Tn10 mutagenesis of *E. coli* SP1.4**

The mini-Tn10 phage vehicle λNK1324 was obtained from the American Type Culture Collection (ATCC number 77347). This mini-Tn10 construct is useful for general transposon mutagenesis in *E. coli* and is marked with a chloramphenicol resistance fragment from Tn9. The procedures for generating transpositions from λ into *E. coli* chromosome were described by Kleckner.[32] λ phage lysate was prepared by propagation of phage in *E. coli* LE392 using the following procedure. Serial dilutions of λ phage stock (0.1 mL, $10^{-1}$ to $10^{-5}$) in TMG buffer (1.2 g/L Tris base, 2.46 g/L $MgSO_4.7H_2O$, 0.1 g/L gelatin, pH 7.4) were prepared in sterile test tubes (13 × 100 mm). An aliquot (0.1 mL, approximately $5 \times 10^8$ cells) of an overnight culture of the *E. coli* LE392 was added to each tube. To each tube 2.5 mL sterile, molten top agar (10 g/L tryptone, 5 g/L NaCl, 7 g/L agar, 45 °C) was added. The contents of each tube were mixed and poured immediately onto a pre-warmed (37 °C) TB1 plate (10 g/L tryptone, 5 g/L NaCl, 11 g/L agar), swirling gently to achieve uniform coverage of the plate. After the agar had solidified, the plates were incubated at 37 °C overnight. A single plaque was picked up with a micropipette and transferred to 10 mL LB plus 10 mM $MgSO_4$ and 0.1 mL of a fresh overnight LB culture of the *E. coli* LE392. After shaking the sample at

37 °C for 4 - 5 h, the culture should gradually become somewhat cloudy, then clear. A few drops of chloroform were added to the culture to make certain that all of the cells had lysed and the culture was then centrifuged (20 000*g*, 10 min, room temperature). Aqueous phage lysate was stored in 1.5 mL microfuge tubes over several drops of $CHCl_3$ at 4 °C.

The titer of the phage lysate was checked by the following way. The phage lysate was diluted $10^{-3}$ and $10^{-6}$ respectively using LB solution. The diluted phage lysate and 0.1 mL *E. coli* LE392 overnight cells was mixed with different ratios and plate in 4 mL top agar on a LB agar plate. The plates were incubated for 8 h at 37 °C. The number of plaque formed was counted and the titer of the phage lysate obtained was normally about $1\text{-}2 \times 10^{10}$ pfu/mL.

Transpositions from λ delivery vehicles were isolated by infecting SP1.4 cells with the phage. *E. coli* SP1.4 was grown overnight on LB medium with 0.2% maltose. Cells were concentrated and resuspended in the LB medium with IPTG (2.5 mM). Cells (0.1 mL, approximately $10^8$-$10^9$ cells) in the concentrated culture were mixed with certain amount of the prepared phage lysate to give a ratio of 0.12 to 0.3 phage per cell. After incubating the cells at room temperature for 30 min, and 37 °C for 90 min, 200 μl 1 M sodium citrate was added. After centrifugation (microfuge, 1 min, room temperature) and resuspension in 200 μl LB with sodium citrate (50 mM), the infected cells were plated on LB plate with chloramphenicol to select for transposition mutants. Normally 100 to 200 colonies per plate were obtained by this procedure. Streaking techniques were then applied to obtain single isolated mutants.[13]

**Y-linker PCR technique**

Construction of the Y-linker was described by Kwon[33] using linker 1 (5'-TTTCTGCTCGAATTCAAGCTTCTAACGATGTACGGGGACACATG) and linker 2 (5'-TGTCCCCGTACATCGTTAGAACTACTCGTACCATCCACAT). Linker 2 (9 μL, 350 ng/μL) was first phosphorylated at the 5' end using T4 polynucleotide kinase (Invitrogen) according to the procedure provided by the manufacturer. After heat denaturation of polynucleotide kinase at 65 °C for 20 min, 9 μL linker 1 (350 ng/μL) was added to a final volume of 29 μL. This mixture was heated at 95°C for 2 min and slowly cooled to room temperature, resulting in the formation of the Y linker at a final concentration of 200 ng/μL.

The genomic DNA was isolated from mini-Tn10 mutants of *E. coli* SP1.4 by the method of Pitcher[17] and completely digested with *Nla*III (New England Biolabs). Approximately 40 ng of the digested DNA was ligated to 1 μg of the Y linker with 1 μl T4 DNA ligase (1 Unit, Invitrogen) in a final volume of 20 μL. After overnight incubation at room temperature, the reaction mixture was diluted with double distilled water to a final volume of 200 μL and heated at 65 °C for 10 min to denature T4 DNA ligase. 4 μL aliquots were used as templates in the PCR amplification. All PCR reactions were performed using a primer specific to mini-Tn10 (5'-TCCCAGAGCCTGATAAAAAC) and a primer specific to Y-linker (5'-CTGCTCGAATTCAAGCTTCT). *Taq* DNA polymerase was used to perform the PCR. The target sequences were amplified through 30 cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min, followed by a final cycle of 72 °C for 10 min. The PCR products were analyzed on a 1.5% agarose gel and purified using Zymoclean Gel DNA

Recovery Kit from Zymo Research Company. Sequences are determined on an ABI PRISM®3100 Genetic Analyzer by biochemistry facility in Michigan State University.

### *E. coli* KL3 derivatives

*E. coli* KL3 was made *tdcA::Tn10*, *ilvI::Tn10*, *narU::Tn10*, *ydaQ::Tn10*, *ksgA::Tn10*, *ykgH::Tn10* by P1 transduction from respective *E. coli* SP1.4 mutants. Colonies were plated onto LB plates containing Cm to select for colonies which had been transduced by mini-Tn10 insertion.

## REFERENCE


1 Pittard, J.; Wallace, B. J. Distribution and Function of Genes Concerned with Aromatic Biosynthesis in *Escherichia coli*. *J. Bacteriol.* **1966**, *91*, 1494-1508.

2 Li, K.; Mikola, M. R.; Draths, K. M.; Worden, R. M.; Frost, J. W. Fed-Batch Fermentor Synthesis of 3-Dehydroshikimic Acid Using Recombinant *Escherichia coli*. *Biotechnol. Bioeng.* **1999**, *64*, 61-73.

3 Knop, D. R.; Draths, K. M.; Chandran, S. S.; Barker, J. L.; von Daeniken, R.; Weber, W.; Frost, J. W. Hydroaromatic Equilibration During Biosynthesis of Shikimic Acid. *J. Am. Chem. Soc.* **2001**, *123*, 10173-10182.

4 Meyer, D.; Schneider-Fresenius, C.; Horlacher, R.; Peist, R.; Boos, W. Molecular Characterization of Glucokinase from *Escherichia coli* K-12. *J. Bacteriol.* **1997**, *179*, 1298-1306.

5 Akowski, J. P.; Bauerle, R. Steady-State Kinetics and Inhibitor Binding of 3-Deoxy-D-*arabino*-heptulonsonate-7-Phosphate Synthase (Trptophan sensitive) from *Escherichia coli*. *Biochemistry* **1997**, *36*, 15817-15822.

6 Furste, J. P.; Pansegrau, W.; Frank, R.; Blocker, H.; Scholz, P.; Bagdasarian, M.; Lanka, E. Molecular Cloning of the Plasmid Rp4 Primase Region in a Multi-Host-Range *tacP* Expression Vector. Gene **1986**, *48*, 119-131.

7 Parker, C.; Barnell, W. O.; Snoep, J. L.; Ingram, L. O.; Conway, T. Characterization of the *Zymomonas mobilis* glucose facilitator gene product (*glf*) in recombinant *Escherichia coli*: Examination of transport mechanism, kinetics and the role of glucokinase in glucose transport. *Mol. Microbiol.* **1995**, *15*, 795-802.

8 Bartolome, B.; Jubete, Y.; Martinez, E.; de la Cruz, F. Construction and Properties of a Family of pACYC184-derived Cloning Vectors Compatible with pBR322 and its Derivatives. *Gene*, **1991**, *102*, 75-78.

9 Balbas, P.; Soberon, X.; Merino, E.; Zurita, M.; Lomeli, H.; Valle, F.; Flores, N.; Bolivar, F. Plasmid vector pBR322 and its special-purpose derivatives - a review. Gene **1986**, *50*, 3-40.

10 Tsang, T.; Copeland, V.; Bowden, G. T. A Set of Cassette Cloning Vectors for Rapid and Versatile Adaptation of Restriction Fragments. *Biotechniques* **1991**, *10*, 330.

11 Farabaugh, M. A. Biocatalytic Production of Aromatics from D-Glucose. M.S. Thesis, Michigan State University, 1996.

12 Chandran, S. S.; Yi, J.; Draths, K. M.; von Daeniken, R.; Weber, W.; Frost, J. W. Phosphoenolpyruvic Acid Availability and the Biosynthesis of Shikimic Acid. *Biotechnol. Prog.* **2003**, *19*, 808-814.

13 Miller, J. H. *Experiments in Molecular Genetics*; Cold Spring Harbor Laboratory: Plainview, NY, 1972.

14 Yi, J.; Li, K.; Draths, K. M.; Frost, J. W. Modulation of Phosphoenolpyruvate Synthase Expression Increases Shikimate Pathway Product Yields in *E. coli*. *Biotechnol. Prog.* **2002**, *18*, 1141-1148.

15 Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Plainview, NY, 1990.

16 Silhavy, T. J.; Berman, M. L.; Enquist, L. W. *Experiments with Gene Fusions*; Cold Spring Harbor Laboratory: Plainview, NY, 1984.

17 Pitcher, D. G.; Saunders, N. A.; Owen, R. J. Rapid Extraction of Bacterial Genomic DNA with Guanidium Thiocyanate. *Letters in Applied Microbiology* **1989**, *8*, 151-156.

18 Miller, J. H. *A short course in bacterial genetics*; Cold Spring Harbor Laboratory: Plainview, NY, 1992.

19 Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72*, 248-254.

20 Schoner, R.; Herrmann, K. M. 3-Deoxy-D-*arabinose*-heptulosonate 7-phosphate synthase. *J. Biol. Chem.* **1976**, *251*, 5440.

21 Gollub, E.; Zalkin, H.; Sprinson, D. B. Assay for 3-deoxy-D-*arabinose*-heptulosonate acid 7-phosphate synthase. *Meth. Enzymol.* **1971**, *17A*, 349.

22 Cooper, R. A. Phosphoenolpyruvate Synthetase. *Methods Enzymol.* **1969**, *13*, 309-314.

23 Fraenkel, D. G.; Horecker, B. L. Pathways of D-Glucose Metabolism in *Salmonella typhimurium*. *J. Biol. Biochem.* **1964**, *239*, 2765-2771.

[24] Paoletti, F.; Williams, J. F.; Horecker, B. L. An enyzmic method for the analysis of D-erythrose 4-phosphate. *Anal. Biochem.* **1979**, *95*, 250-253.

25 (a) Tao, H.; Gonzalez, R.; Martinez, A.; Rodriguez, M.; Ingram, L. O.; Preston, J. F.; Shanmugam, K. T. Engineering a homo-ethanol pathway in *Escherichia coli*:

increased glycolytic flux and levels of expression of glycolytic genes during xylose fermentation. *J. Bacteriol.* **2001**, *183*, 2979-2988. (b) Gonzalez, R.; Tao, H.; Purvis, J. E.; York, S. W.; Shanmugam, K. T.; Ingram, L. O. Gene Array-Based Identification of Changes That Contribute to Ethanol Tolerance in Ethanologenic *Escherichia coli*: Comparison of KO11 (Parent) to LY01 (Resistant Mutant) *Biotechnol. Prog.* **2003,** *19*, 612-623.

26 Hocking, R. R. *Methods and Applications of Linear Models. Regression and the Analysis of Variance.* 1996, New York: Wiley.

27 Eisen, M. B.; Spellman, P. T.; Brown, P. O.; Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14863-14868.

28 Cadwell, R. C.; Joyce, G. F. Mutagenic PCR. *PCR Methods Applic.* **1994**, *4*, S136-S139.

29 Cheng, S.; Fockler, C.; Barnes, W.; Higuchi, R. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci.* **1994**, *91*, 5695-5699.

30 Ebbighausen, H.; Weil, B.; Kraemer, R. Isoleucine Excretion in *Corynebacterium glutamicum*: Evidence for a Specific Efflux Carrier System. *Appl. Microbiol. Biotechnol.* **1989**, *31*, 184-190.

31 Cromartie, T. H.; Polge, N. D. In *U.S.*; (Syngenta Limited, UK). Method of Detecting Shikimic Acid. Int. Patent WO00/49399, August 24, 2000.

32 Kleckner, N.; Bender, J.; Gottesman, S. Uses of Transposons with Emphasis on Tn10. *Methods in Enzymology* **1991**, *204*, 139-180.

33 Kwon, Y. M.; Ricke, S. C. Efficient Amplification of Multiple Transposon-Flanking Sequences. *J. Microbiol. Methods* **2000**, *41*, 195-199.